A METHODOLOGY TO ESTABLISH THE CRITICALITY OF ATTRIBUTES IN OPERATIONAL TESTS

A THESIS

Presented to

The Faculty of the Division of

Graduate Studies

By

Gary Steven Williams

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in Operations Research

Georgia Institute of Technology

October, 1975

A METHODOLOGY TO ESTABLISH

THE CRITICALITY OF ATTRIBUTES IN OPERATIONAL TESTS

Approved:

Leslie G. Callahan, Chairman

Harrison M. Wadsworth

Paul S. Jones

Date approved by Chairman: 31 Oct 1975

ACKNOWLEDGEMENTS

I would like to express my appreciation to the following key individuals who contributed in the preparation of this thesis:

Dr. Leslie G. Callahan, my thesis advisor, has been the principal source of support and encouragement from inception to completion of this endeavor.

Dr. H. M. Wadsworth and Dr. P. S. Jones, members of my reading committee, have provided many constructive comments and much support which have considerably enhanced this effort.

The United States Army Operational Test and Evaluation Agency has provided financial support and technical expertise which helped greatly.

Finally to my wife and son, Ellen and Chris, I owe the most for the sacrifices they made in order that this thesis could be completed.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF ILLUSTRATIONS

SUMMARY

Operational testing is an integral part of the materiel acquisition cycle in the Army procurement process. It is oriented towards the evaluation of a developmental item under realistic conditions as part of an actual troop unit. The test design phase is an essential element of operational testing.

In order to facilitate the selection of measures of effectiveness used in these tests, the critical attributes which "best" discriminate between acceptable and unacceptable systems or subsystems need to be identified. This thesis addresses a method which provides a basis for the selection of these critical attributes. Once these attributes are identified, the test designers of subsequent operational tests may use this information to assist them in the test design phase.

The current test structure in operational testing is not ameanable to the standard application of multivariate statistics. There is only one replication of each test of these large systems and the data collection procedure precludes direct determination of relationships among the attributes. Consequently, the methodology developed in this thesis encompasses a means to combine results from past tests with subjective information to determine the relationship, in terms of covariances, between each two attributes. This information is incorporated with subjectively obtained acceptable and unacceptable mean vectors in stepwise discriminant analysis.

It is concluded that multivariate analysis techniques may be a valuable aid in determining which attributes contribute more in distinguishing between successful and unsuccessful systems. It is also concluded that the current test design for operational testing can be modified to facilitate a broader use of multivariate statistical analysis techniques. This modification should permit (1) the correlations among the attributes to be objectively determined and (2) the marginal normality of observations for each attribute to be validated.

CHAPTER I

INTRODUCTION

Background of the Problem

The United States Army Operational Test and Evaluation Agency (OTEA) is a Department of the Army Field Agency under the Army Chief of Staff. OTEA's mission is to support the materiel acquisition and force development processes by (1) exercising responsibility for all operational testing (OT), (2) managing force development testing and experimentation (FDTE), and (3) managing joint user testing for the army. It must insure that user testing is effectively planned, conducted, and evaluated with emphasis on adequacy, quality, and credibility. It actively participates in the conduct of and provides independent evaluations of operationl tests conducted on major and selected nonmajor systems, as well as major FDTE and other systems designated by appropriate authority (38).

Operational testing is an integral part of the materiel acquisition cycle. It is oriented towards the evaluation of a developmental item under realistic conditions as part of an actual troop unit. The purposes of operational testing are: (1) evaluation of the item's desirability compared to equipment already in the inventory; (2) evaluation of military utility, operational effectiveness, and operational suitability; (3) assessment of the need for modification; and (4) assessment of the adequacy of organization, doctrine, and tactics (34).

Force development testing and experimentation involves troop tests, field tests, and experiments performed by or for the users. The tests support the force development process by examining the impact, potential, or effectiveness of selected concepts, doctrine, organization, and materiel (34). A test can support the materiel acquisition process by providing data to assist in the establishment of the required operatonal capability, to develop fundamental data necessary for a full understanding of the performance of a materiel system, or to assist in validating doctrine and tactics to counter threat response to a system once deployed (38).

The full development of an operational test from the initial planning phase through the final test report is a long and detailed process. This process delineates exactly how each aspect of conducting the tests is developed. However, the aspect of prime consideration in this paper is the development of measurable attributes, i.e., measures of effectiveness (MOE), of the tests.

The first step in the developmental process is the initial approach, that is, the listing of tentative operational issues. These issues are the aspects of the system's capability that must be questioned before the system's effectiveness is known. They are broad in nature and are not necessarily directly measurable. These tentative operational issues are evaluated and critical operational issues are developed on the basis of relevance, importance, and risk. Finally, the critical issues are consolidated as necessary, and the issues for operational testing are selected after considering their validity, practicality, and relative costs.

When the operational issues have been refined, statements of test objectives are developed. These statements identify the evidence required to address particular issues. The test objectives may still not be measurable; hence, they must be divided into subobjectives. The subobjectives are further subdivided into lower levels of data requirements until measurable requirements emerge. A data requirement is finally in a form suitable for measurement when it can be answered by a number.

These data requirements, MOE, are later refined in the Final Test Design. The refinement is necessary because one aspect of effectiveness often has several possible measures and not all are needed. The questions of redundancy and duplication are addressed, the advantages and disadvantages are considered, and a decision is made as to which measures to employ in the tests. The selection of measures involves some risk of selecting inferior measures; consequently, some special assistance is needed in making this selection.

This special assistance can come in the form of specialists who are familiar with doctrine, organization, human factors, logistics, and threat, and in the form of information developed from previous operational tests.

> Normally, operational testing considerations begin with the development of the test item and conclude with the publication of the final report. However, it should be noted that one of the paths to improved operational test methodology begins after the final report is published. That is, data analyzed in the post-test-report period as a means to further refine the nature of influencing factors can be used to improve the state of the art for subsequent operational testing. Thus, field testing not only contributes to system evaluation objectives, but has the potential for contributing to all future operational tests. This final contribution may be quite as important in the long run as answering the test objectives (38).

Definition of the Problem

A key area of current interest to OTEA is the evaluation of tactical command and control systems. One accepted definition of a tactical command and control system is an arrangement of personnel, facilities, and the means for information acquisition, processing, and dissemination employed by a commander in planning, directing, and controlling tactical operations.

The introduction of sophisticated computer-based command and control systems into the materiel acquisition process raises a problem in the operational testing of such systems. In the past, operational tests have been able to evaluate hardware and software independently; however, there is presently a need to evaluate the operational effectiveness of the entire system. This system consists of the hardware, software, and personnel interface under complex operational conditions. The evaluation should take into account the interplay of all relevant influencing variables.

The general problem of this thesis is a current requirement for the development of a detailed methodology for designing, planning, and evaluating the results of operational tests and evaluations of complex command and control systems. The specific problem of this thesis impacts directly on the design, plan, and evaluation of operational tests. This problem is: How can the size of these complex tests be reduced without reducing the amount of information about the system being tested?

Purpose of this Thesis

The purpose of this thesis is to develop a methodology which will provide a rational basis for selecting critical attributes of complex command and control systems. This selection process will permit the deletion of the evaluation of non-critical attributes from subsequent operational tests.

Review of Literature

The practical use of multivariate statistics to determine the criticality of attributes in operational tests is virtually untrodden ground. Both computer based and manual literature searches revealed no direct references to this area of study. There are, however, many articles and books about multivariate statistics. The multivariate techniques of principal components analysis (3, 4, 9, 15, 30), factor analysis (3, 4, 6, 9, 15, 30), discriminant analysis (3, 4, 9, 17, 18, 19, 21, 30), and cluster analysis (1, 5, 7, 16, 20, 23, 24, 5, 32, 36, 40) were considered as possible techniques to be utilized. All of these techniques were eliminated from consideration because of the design of the tests. The methodology, as will be seen later, does not provide for any replications of the test. The operational tests are so complex that the extremely high costs preclude replications. Multiple classification analysis, the Automatic Interaction Detection System, was considered, but was eliminated because of the requirement for a large sample of data (38).

Interviews and extended discussions with representatives of the Methodology Branch, Test Design Division, Operational Test and Evaluation Agency and the Test and Evaluation Division, Command Control and Communications Directorate, Modern Army Selected Systems Evaluation and Review resulted in valuable insight into the problem area. However, techniques that could be modified or utilized directly were not available from these sources. The relative lack of prior study in this field and the nature of the tests themselves led to the development of the methodology presented in Chapter III.

General Approach and Overview

The multidimensional aspect of complex systems lends itself to the application of multivariate analysis. The complexity of these systems inherently causes the tests that are used to evaluate the systems to become large and unwieldy. Chapter II will address the methodology of these tests. It will emphasize the design and evaluation procedure for the tests.

Chapter III will review the analysis leading to the selection of the methodology. The detailed procedure for each step of the methodology to establish the criticality of attributes will then be presented in Chapter IV.

An integral part of this methodology is the use of a computer to facilitate the manipulation of the multivariate data. The computer programs used will be discussed in Chapter IV in the sequence in which they are utilized.

A demonstration of the methodology will be presented in Chapter V in order to illustrate the entire procedure. The final conclusions and recommendations will be presented in Chapter VI.

CHAPTER II

EXISTING TEST STRUCTURE

Methodology of the Tests

To determine critical attributes of operational tests, it is necessary to understand the intricacies of those tests. Each test is different from every other test in that the attributes which are measured vary according to the operational issues involved. Although specific tests differ, the methodology by which those tests are evolved from operational issues to conclusion is similar for a large set of tests.

Normally, OTEA provides a Final Test Design Plan for each test prior to the start of detailed planning at the test site. Exceptions occur for tests that are executed for OTEA by Modern Army Selected Systems Evaluation and Review (MASSTER). In this latter case, the Final Test Design Plan is prepared by MASSTER for OTEA approval. The test structure examined in this chapter illustrates the general methodology. A specific test is used as an example in order to concisely illustrate the methodology.

The Division Command Post Test (Test Number FM 286) was selected as the example. Its purpose is to evaluate a proposed division command post (CP) system and thus is a test of a command and control system. The methodology presented in FM 286 is representative of the general methodology used by OTEA. The data from this test are not classified and thus are available for analysis. Because the test was conducted in January

1975, some of the principles involved in the test are available to explain unanswered questions that are not covered in the written plans and reports. The test structure presented in this chapter is derived from (1) the Detail Plan for Execution (FM 286) (10), (2) the Division Command Post Test Report (FM 286) (11), (3) the MASSTER Test Officer's Planning Manual (34), and (4) interviews with representatives of MASSTER. In the interest of clarity, the administrative details of the test will not be related, but they are available in the references mentioned above for the interested reader.

Pattern of Analysis

The purpose of the test FM 286 is to evaluate a proposed division command post (CP) system. The results of the test are to be utilized to support recommedations concerning tactical organization, equipment, and command and control doctrine and procedures. They will be the basis for subsequent changes to Tables of Organization and Equipment.

The objectives that were derived from the operational issues mentioned in Chapter I are the evaluation of the efficiency of the command post in command and control of division tactical operations and the evaluation of the vulnerability of the command post during division tactical operations. Efficiency is defined as "a measure of the degree to which a system performs a set of defined tasks or mission requirements." (10) Vulnerability is defined as "a measure of the susceptibility of the command and control system to any reasonable means through which its combat effectiveness might be reduced." (10) Note that the system is being evaluated and not the performance of the players within the system.

For the purpose of continuity, the development of objective 1, efficiency, is pursued and the development of objective 2, vulnerability, is omitted.

Army Regulation 310-25 defines a system as "an integrated relationship of components aligned to establish functional continuity toward the successful performance of a defined task or tasks." Therefore, the division command and control system was divided into the subsystems of command, operations, intelligence, and combat service support.

The primary functions of the command subsystem are the management functions of planning, organizing, directing, coordinating, and controlling. In order to find a measure of efficiency for the command subsystem, a measure of efficiency for each of the functions listed above must be found. Thus the management functions of the command subsystem become data requirements. The data requirements are not measurable. Consequently, they must be divided and subdivided until measurable data requirements are developed. An abbreviated pattern developed in this manner is shown in Figure 1.

The primary staff functions are preparing plans, orders, and reports; providing information; and supervising the execution of plans and orders. Considering operations, intelligence, and combat service support as primary staff functions, each subsystem is subdivided according to these functions. Since all subsystems have the same function, only the operations subsystem will be pursued. As in the command subsystem, the functions of the operations subsystem are unmeasurable data requirements that must be further subdivided until measurable data requirements are developed. An abbreviated pattern developed in this manner is shown in Figure 2. The intelligence and combat service support subsystems are

- 1.1. Comd Gp Efficiency
  - 1.1.1 Comd Gp Planning
    - 1.1.1.1. Effect of TOE on Planning
    - 1.1.1.2. Effect of Org Struc on Planning
  - 1.1.2 Comd Gp Organizing
    - 1.1.2.1. Effect of TOE on Organizing
    - 1.1.2.2. Effect of Org Struc on Organizing
    - 1.1.2.3. Effect of SOP on Organizing
  - 1.1.3 Comd Gp Directing
    - 1.1.3.1. Effect of TOE on Directing
    - 1.1.3.2. Effect of Org Struc on Directing
    - 1.1.3.3. Effect of SOP on Directing
  - 1.1.4 Comd Gp Coordinating
    - 1.1.4.1. Effect of TOE on Coordinating
    - 1.1.4.2. Effect of Org Struc on Coordinating
    - 1.1.4.3. Effect of SOP on Coordinating
  - 1.1.5 Comd Gp Controlling
    - 1.1.5.1. Effect of TOE on Controlling
    - 1.1.5.2. Effect of Org Struc on Controlling
    - 1.1.5.3. Effect of SOP on Controlling
  - 1.1.6 Other Events that Affect Efficiency

Figure 1. Abbreviated Pattern of Analysis - Objective 1.

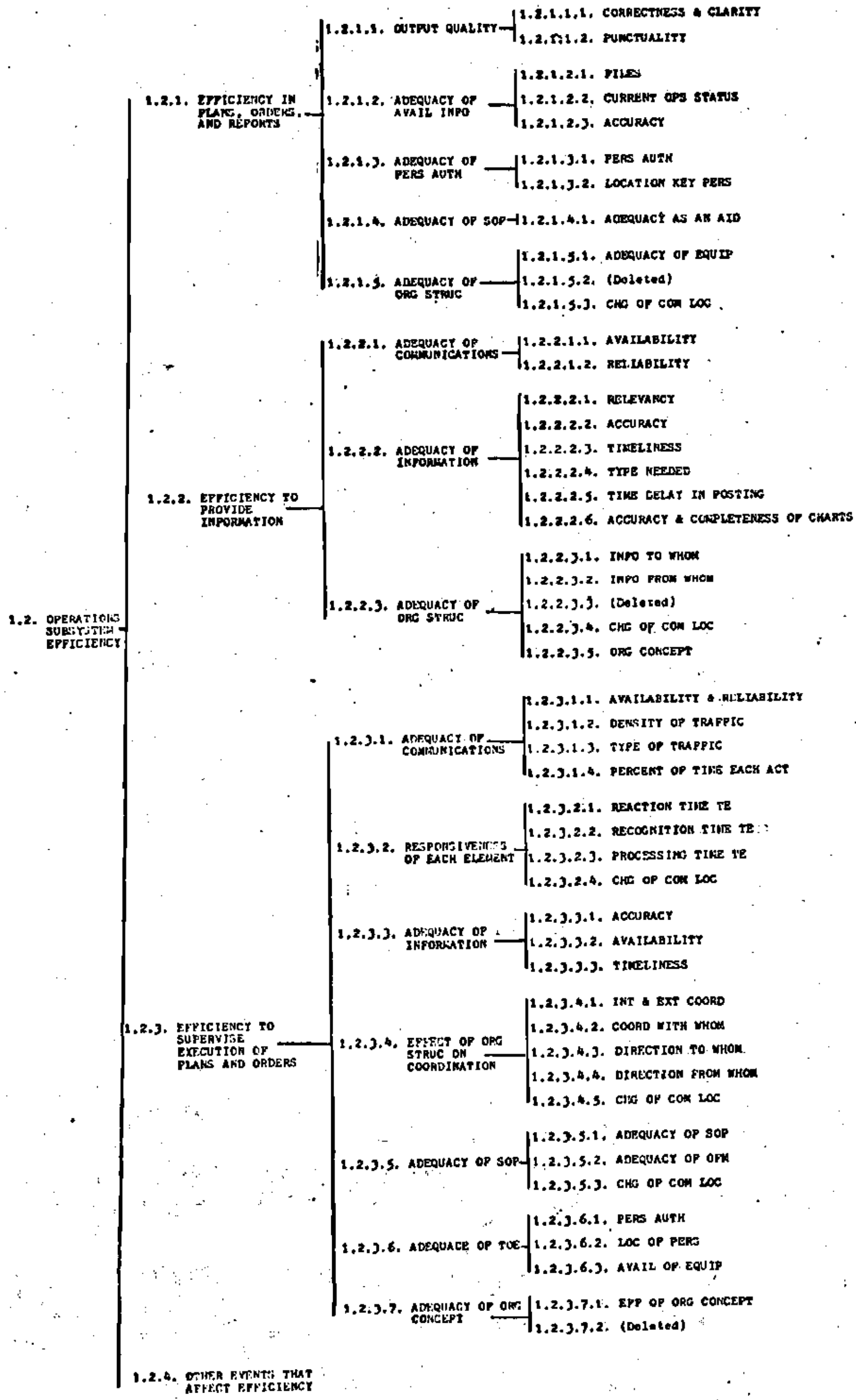


Figure 2. Abbreviated Pattern of Analysis, Operations.

analogous to that for operations.

The measures of effectiveness (MOE) that were used to develop the pattern of analysis are a set of dependent or response variables which demonstrate the adequacy of the command post system to accomplish mission requirements under specific conditions. These specific conditions are the independent variables.

Evaluation Plan

The evaluation effort for objective 1 was essentially subjective. It was reinforced with quantitative data where feasible. This subjective evaluation was based primarily on the end-of-test reports prepared by the evaluators and players. Extensive use of rating type questions, using semantic differentially scaled responses, were used in the daily questionnaire. These ratings allowed the opinions of the players and evaluators to be quantified on a daily basis and subsequently aggregated within each major staff section. When aggregated, the daily ratings also provided the test analyst with numerical performance indicators at each level of the analysis. Objective measurements, such as accuracy of maps and charts, are also included in the assessment of dependent variables. All of these types of input provide a means to assess the adequacy of the command post concept to accomplish mission requirements.

A subjective extension of the player and evaluator semantic differential responses is an adequate/inadequate evaluation scheme. This is the basis for the development of criteria used by the test analyst to assess the measure of performance (MOP) for the command post concept. Criteria for the development of conclusions for objective 1 are defined in the following manner:

Adequate- Combined ratings and rationale demonstrated that the section could accomplish functions or mission requirements under test conditions.

Borderline- Although ratings and rationale may have demonstrated that the section could accomplish the functions or requirements, there were negative ratings and rationale that detracted from the overall ability to accomplish mission requirements. These negative factors had to be correctable without major organizational or functional changes to the OFM or SOP.

Inadequate- Ratings and rationale demonstrated that the section could not accomplish mission requirements. Major organizational or functional changes had to be made to either the OFM or SOP (11).

OFM and SOP are acronyms for organizations and functions manual and standing operating procedures, respectively.

To enhance the evaluation plan, objective 1 is functionally divided into areas for which findings are generated, Figure 3. Note that this functional division breaks down each subsystem into a section that is operationally responsible for the subsystem. The elements within that section are designated. These elements comprise the level at which the measures of performance are evaluated.

The data collection plan involves both players and evaluators in obtaining data on MOE's presented in the plan of analysis. The evaluators are screened to insure their qualifications and to insure the credibility of their observations and evaluations based on grade, military occupational speciality (MOS), command and staff experience, military schooling, etc. There are three categories of evaluators. Category 1 consists of officers in the grades of 05 and 06. They are assigned to the subsystem-section level to evaluate the effect of the organizational concept on staff and section performance by observation

SYSTEM	SUBSYSTEM	SECTION	ELEMENT	FUNCTIONS
EFFICIENCY OF CP SYSTEM	OPERATIONS	G3	OPS	(See 1,2,3)
			FS	
			TACP-F	
			PLANS	
			FSE	
			DAME	
			LNO	
	INTELLIGENCE	G2	CM&D	(See 1,2,3)
			A&P	
			SSE/EWE	
			R&S	
			TACP	
			P&A	
			CI&I	
	CSS	G1&G4	G1 (Main)	(See 1,2,3)
			G1 (DSA)	
			G4 (Main)	
			G4 (DSA)	

1. Prepare plans, orders, and reports
2. Provide information
3. Supervise execution of plans and orders

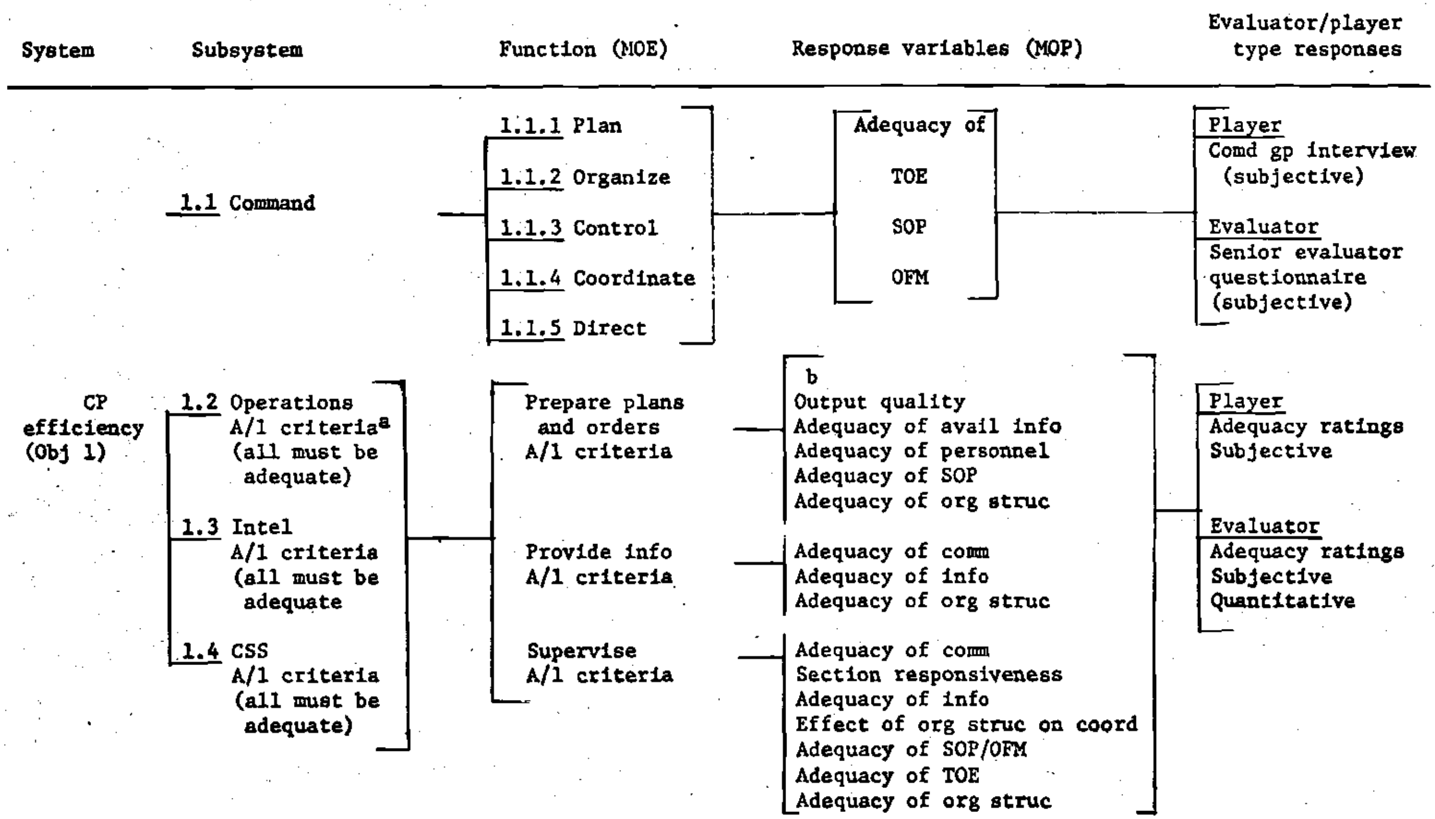
Figure 3. Functional Areas Used in Evaluation, Objective 1.

and review of staff outputs. Category 2 evaluators consist of officers in the grades of 03 and 04. They are assigned to the element level to evaluate the performance of one or more selected staff elements and to record staff performance as appropriate. Category 3 consists of enlisted men in the grades of E6 and E7. They are assigned to the element level to collect data from selected maps and charts, prepare CP layout sketches, inventory major items of equipment, and record displacement data.

There are seventeen types of questionnaires and data forms that are used by the players and evaluators. The questionnaires and data forms are tailored as to the data source, frequency of submission, and level of required detail. These are completed and submitted according to the Data Collection Plan, Appendix E to the DPE.

Corresponding to the categories of evaluators are players that complete, respectively, the same questionnaires and data forms. Thus, there is a dual rating system. The data from the players and evaluators is assimilated according to the Data Reduction Plan, Appendix F of the DPE for FM 286. There are many details in the reduction plan; however, the salient feature is the method by which player and evaluator observations are combined in order to formulate an evaluation. Figure 4 presents an overview of the interaction of player-evaluator response.

Prior to a discussion on data reduction, it is necessary to explain the acronym EEA. EEA, essential elements of analysis, are those data requirements that have been developed for a specific test. The first level EEA corresponds to the subsystems; the second level EEA corresponds to the functions; and the remaining levels of EEA correspond respectively to the succeeding subdivisions.



[a] Adequate/inadequate criteria [b] Adequate/borderline/inadequate criteria

Figure 4. Evaluation Methodology Concept for Objective 1.

The ratings for the third level EEA (e.g., 1.2.2.2, Adequacy of Information) are basically obtained in the following steps:

Step 1: Ratings of fourth level EEA are made by each category 2 and 3 evaluator and player for each element. These ratings are based on a five-point adjectival rating scale where 1 is considered best and 5 is considered worst (e.g., 1 corresponds to adequate, 3 corresponds to borderline, and 5 corresponds to inadequate). The frequency of observations that fall in the five rating categories for each element and data requirement are recorded separately for both evaluators and players. If an observation is placed in rating category 1, it receives a weighted value of 1. If an observation is placed in rating category 2, it receives a weighted value of 2. This procedure extends analogously for the remaining three rating categories. If an observation is not made, a sixth rating is available. This rating is considered as having no weight and does not affect the evaluation procedure (See Table 1).

Step 2: The fourth level ratings are summed across the rating categories for each element of the data requirement. The frequencies of the rating scores are multiplied by their respective weights and an element score is obtained by dividing the sum of these by the total number of observations by that element. This is done for both players and evaluators at this level (See Table 2).

Step 3: A third level EEA score is obtained for each element by taking a grand average across the fourth level EEA (data requirements). This is done for players and evaluators (See Table 3).

Table 1. Frequency of Observations

Data Requirement 1.2.2.2.2.

PLAYER RATINGS					
ELEMENT	1	2	3	4	5
OPS	10	23	4	0	0
FS	0	9	0	1	0
TACP	1	3	1	0	0
PLANS	2	13	0	0	0
FSE	15	18	2	0	0
DAME	4	8	8	0	0
LNO	5	0	0	0	0

EVALUATOR RATINGS					
ELEMENT	1	2	3	4	5
OPS	6	4	0	0	0
FS	8	1	1	0	0
TACP	2	7	0	1	0
PLANS	2	6	2	0	0
FSE	4	6	0	0	0
DAME	1	7	0	2	0
LNO	0	0	0	0	0

Table 2. Fourth Level Element Scores

Data Requirement 1.2.2.2.2.

PLAYER RATINGS						
ELEMENT	1	2	3	4	5	ELEMENT SCORE/ No. of OBSER.
OPS	10	23	4	0	0	1.84/37
FS	0	9	0	1	0	2.20/10
TACP	1	3	1	0	0	2.00/50
PLANS	2	13	0	0	0	1.87/15
FSE	15	18	2	0	0	1.48/25
DAME	4	8	8	0	0	2.20/20
LNO	5	0	0	0	0	1.00/50

Example of Element Score for OPS:

$$\frac{10(1) + 23(2) + 4(3) + 0)4) + 0(5)}{37} = 1.84$$

EVALUATOR RATINGS						
ELEMENT	1	2	3	4	5	ELEMENT SCORE/ No. of OBSER.
OPS	6	4	0	0	0	1.40/10
FS	8	1	1	0	0	1.30/10
TACP	2	7	0	1	0	2.00/10
PLANS	2	6	2	0	0	2.00/10
FSE	4	6	0	0	0	1.60/10
DAME	1	7	0	2	0	2.30/10
LNO	0	0	0	0	0	2.30/10

Example of Element Score for OPS:

$$\frac{6(1) + 4(2) + 0(3) + 0(4) + 0(5)}{10} = 1.4$$

Table 3. Third Level Element Score
(Category 2 and 3)

Data Requirement 1.2.2.2.

PLAYER ELEMENT SCORES/NUMBER OF OBSERVATIONS				
ELEMENT	1.2.2.2.1	1.2.2.2.2	1.2.2.2.3	1.2.2.2.
OPS	1.78/37	1.84/37	2.36/36	1.99/110
FS	2.10/10	2.20/10	3.00/10	2.43/30
TACP	1.20/50	2.00/50	4.40/5	2.53/15
PLANS	1.60/15	1.87/15	4.07/15	2.51/45
FSE	1.30/20	1.48/25	3.40/20	2.01/65
DAME	1.72/25	2.20/20	4.36/25	2.80/70
LNO	1.00/50	1.00/50	2.60/50	1.53/15

Example of Third Level Element Score for OPS:

$$\frac{1.78(37) + 1.84(37) + 2.36(36)}{110} = 1.99$$

EVALUATOR ELEMENT SCORES/NUMBER OF OBSERVATIONS				
ELEMENT	1.2.2.2.1	1.2.2.2.2	1.2.2.2.3	1.2.2.2.
OPS	1.40/10	1.40/10	1.80/10	1.53/30
FS	1.30/10	1.30/10	2.10/10	1.57/30
TACP	1.80/10	2.00/10	3.40/10	2.40/30
PLANS	1.70/10	2.00/10	3.90/10	2.53/30
FSE	1.40/10	1.60/10	3.10/10	2.03/30
DAME	1.90/10	2.30/10	4.60/10	2.93/30
LNO	-	-	-	-

Example of Third Level Element Score for OPS:

$$\frac{1.4(10) + 1.4(10) + 1.8(10)}{30} = 1.53$$

Step 4: Category 1 players and evaluators make independent observations on the third level EEA. A simple average is obtained for each (See Table 4).

Step 5: A grand average is taken across the elements of third level EEA scores to obtain a third level EEA score. This is done for both players and evaluators. D. R. 1.2.2.2.

Player

$$\frac{1.99(110)+2.43(30)+2.53(15)+2.51(45)+2.01(65)+2.80(70)+1.53(15)}{350} = 2.26$$

Evaluator

$$\frac{1.53(30)+1.57(30)+2.4(30)+2.53(30)+2.03(30)+2.93(30)}{180} = 1.91$$

Step 6: A mean is then obtained of the score for the category 2 and 3 players (Step 5) and the score for the category 1 player (Step 4).

D. R. 1.2.2.2.

$$\frac{2.26+3.4}{2} = 2.83$$

A score for the evaluators is obtained in a similar manner.

$$\frac{1.91+2.0}{2} = 1.96$$

Table 4. Third Level Element Score
(Category 1)

Data Requirement 1.2.2.2.

PLAYER RATINGS					
ELEMENT	1	2	3	4	5
G3	0	2	1	0	2

Category 1 Third Level Score:

$$\frac{0(1) + 2(2) + 1(3) + 0(4) + 2(5)}{5} = 3.4$$

EVALUATOR RATINGS					
ELEMENT	1	2	3	4	5
G3	1	1	2	1	0

Category 1 Third Level Score:

$$\frac{1(1) + 1(2) + 2(3) + 1(4) + 0(5)}{5} = 2.0$$

Step 7: The mean of the player and evaluator scores is the overall rating for the third level EEA.

$$\frac{2.83+1.96}{2} = 2.39$$

Once all of the overall scores for the essential elements of analysis are obtained, an independent evaluator considers all input, subjective and objective, and formulates a final rating for each third level EEA. The minimum acceptable standard at this level is "borderline." Ratings of acceptable, borderline, or unacceptable are established for each section. Each of the staff sections is then subjectively evaluated at progressively higher levels of EEA as adequate or inadequate in the performance of its primary functions. A rating of adequate is the minimum acceptable standard at higher levels of evaluation.

Analysis of the Test Structure

The purpose of this analysis is not to evaluate the methodology underlying the structure of the test. Rather, the purpose is to examine the salient features of the test in order to better understand the relationship of the data requirements to the test structure.

The nature of the test is highly subjective. The objective inputs do assist in the final evaluation; however, the subjective inputs obtained by the use of the semantic differentially scaled responses and general observations have a greater impact upon the final evaluation of the system. The objective inputs and subjective inputs using the five-point adjectival scale do provide a profile of the individual measurable data requirements. They provide a frequency distribution of the number of times the data requirements were observed in a particular status, i.e., the number of observations that fell in the rating categories 1 through 5. The manner in which the observations were taken precludes the determination of relationships between data requirements (correlation). The number of total observations and time of observations varied greatly. The range of total observations per data requirement is illustrated by the data used in the demonstration of Chapter 4.

The intervals of the rating scale restrict the range of observations. If an observer evaluates a particular measurable data requirement as slightly above borderline, but not totally adequate, he has only one alternative. That alternative is to assign a rating of 2. The rating of 2 reflects that the data requirement is evaluated exactly halfway between borderline and adequate. This indicates that the range of the rating scale might need to be extended. The optimal number of intervals between which an observer can discriminate in making an evaluation is an open question for study.

The number of measurable data requirements in this type of test is quite large. Since multiple observations are being made for each measurable data requirement, the size of the problem, computationally, can easily get out of control. The level at which the data requirements are measurable also varies. Thus there is the added problem of comparing data requirements from different level EEA.

The evaluation procedure is based upon the elements at each level rather than the measurable data requirements. As seen earlier, evaluations at each level EEA are made that reflect the status of the elements.

The observations of the data requirements are not carried forward. This is compounded by the means by which the overall evaluations are computed. The averaging process dilutes the rating category 5 observations. By the time the final third level EEA are calculated, the "worse" evaluations are diluted. This situation indicates a need for a method to consider all observations for each data requirement.

CHAPTER III

REVIEW OF POTENTIAL MULTIVARIATE ANALYSIS TECHNIQUES

Introduction

The purpose of this thesis, as stated in Chapter I, is to develop a methodology which will provide a rational basis for the selection of critical attributes of complex command and control systems. Chapter II presented an actual test that illustrates the methodology that is currently used to evaluate complex command and control systems. In this chapter the term "critical attributes" will be defined and potential multivariate analysis techniques for determining critical attributes will be evaluated.

The attributes of a system may be defined in several ways. They may be an integral part of the evaluation procedure used in a particular test. In the test discussed above, the evaluation procedure included observing the measurable data requirements, formulating a value for each element, and carrying the element evaluation up the levels of EEA in order to obtain section evaluations. Here the elements could be considered to be the attributes. In that procedure, the observations made on the specific data requirements were lost in the evaluation process.

The attributes may also be defined as the measurable data requirements of the tests. The evaluation procedure may vary between two tests; however, the methodology by which operational issues are subdivided into measurable data requirements is stable. If the measurable data requirements are considered as the attributes, all observations have a visible impact on the overall resulting evaluation up to the third level EEA. In order to insure the greatest degree of applicability for the methodology presented here, the term "attribute" will be construed to mean measurable data requirements.

The critical attributes, or critical measurable data requirements, of a system are those attributes which impart maximum information to its evaluation. In the final analysis, the evaluation of a system classifies the system as either acceptable or unacceptable. Thus, the attributes which contribute most to deciding whether a system is acceptable or unacceptable are classified as critical and those which contribute least are classified as noncritical. The degree of criticality or noncriticality is dependent upon a given situation, or in this case, upon a given operational test. As will be seen later in this chapter, the degree of criticality can be controlled depending on the parameters used in the methodology presented herein.

In developing a methodology to determine critical attributes, the practical use of the methodology was the foremost consideration. The extreme costs of operational tests preclude replications and hence preclude the standard multivariate analysis techniques such as principal component analysis, factor analysis, discriminant analysis, and cluster analysis. The procedure developed had to take into consideration the test design process, the structure of the test, and the form of the data.

The central theme of the design process was the development of measurable data requirements from the initial operational issues. In the design phase, once all measurable data requirements (attributes) were formulated, a subjective process delineated those attributes considered necessary. This process took into consideration the possible necessity of obtaining redundant information by taking manual measures to back up sophisticated instrumentation. It also took into account the need for duplication of measures. This includes observing two attributes that measure the same thing. The great degree of work and expertise that went into making the decisions of the initial attributes should be used to the maximum extent possible.

The structure of the test should be considered. The dendritic breaks down the data requirements into distinct, yet related, levels of essential elements. In FM 286, (see Figure 2), the operations subsystem was broken down into three major areas of consideration: preparing plans, orders, and reports; providing information; and supervising plans and orders. These in turn were subdivided. The possible advantageous use of the structure should be considered in the development of a methodology.

The form of the data should also be considered. Initially, the data is compiled in the form of frequency distributions. If the data is used at the attribute level, a more accurate picture can be obtained of the attributes under consideration.

The methodology presented here evolved from a detailed study of FM 286, the considerations given above, and review of multivariate analysis theory and associated techniques. The general methodology will be presented in the following section. A more detailed discussion will then be presented that explains the procedures and techniques in greater detail.

Applicability of Multivariate Techniques

Considering the definition of critical attributes and the delineation of the criteria for selecting a methodology that were derived from the existing test structure, the task of finding a methodology for selecting critical attributes of a system is somewhat facilitated. The form of the data on the attributes is a frequency distribution with observations on a range of discrete numbers from 1 to 5. There is only one replication of the test and the time interval for reporting observations is so large that statistical correlations cannot be obtained. Thus, there are three alternatives open for investigation.

A totally subjective methodology can be developed that involves no statistical inference. This alternative is considered infeasible because of the nature of the problem. The number of variables is so great that the time involved would extend the design phase past reasonable suspense dates. The number of people necessary for a totally subjective method could easily exceed the number available. Here, it must be understood that the design phase of one test may be, and probably is, conducted simultaneously with a number of other tests. An additional factor is the fact that the degree of criticality is difficult to define subjectively. For these reasons, the totally subjective approach was eliminated from consideration.

A second approach might be totally objective, represented by traditional multivariate statistical techniques. Since there is only one replication of the test, initiating a purely statistical approach would require the simulation of a multiple set of data. The generation of the data would have to be accomplished by generating sets of independent data for each attribute. This would have to be done because a covariance matrix and mean vector are necessary to generate a multivariate distribution. The sets of independently generated data for the attributes would not necessarily produce reliable statistical inference.

The third approach is the use of subjective input in conjunction with multivariate analysis techniques. The objection to the totally objective approach was the lack of a covariance matrix and mean vector from which the generation of multivariate observations could be facilitated. If a covariance matrix and mean vector can be developed using the available data and subjective information, then the generation of multivariate observations is feasible.

At this point in the development of the methodology, it is assumed that a mean vector and covariance matrix can be subjectively produced. Having made this assumption, the available multivariate analysis techniques will be reviewed in order to find a technique that can be utilized to distinguish critical attributes.

In selecting an appropriate technique applicable in the context of operational tests, there are two constraints that restrict the flexibility of technique selection. The first constraint is that a technique is sought whereby the attributes that contribute least to the evaluation of the system are delineated. This is equivalent to specifying those attributes that contribute the most to the evaluation. A second constraint is that the critical attributes are those attributes which best establish whether a system is acceptable or unacceptable. Thus, the technique should involve distinguishing between acceptable and unacceptable populations.

The first constraint reduces the field of multivariate analysis techniques to four areas: principal component analysis, factor analysis, cluster analysis, and discriminant analysis. Principal component analysis, largely attributed to Hotelling (4), deals with the coordinate structure of multivariate observations. This technique seeks to make linear combinations of the variables (principal components) such that each of the linear combinations captures as much variation in the vector of variables as possible. At the same time, each principal component is formulated so that it is linearly independent of all the other principal components. Those principal components that contribute most to the variance would then be the critical variables. This technique is rejected for basically two reasons. The first reason is that the definition of the term "critical variable" does not coincide with the given definition of critical attribute. The second is that principal component analysis deals with one set of variables and one population. The desired technique must involve two populations.

Factor analysis is an extension of principal component analysis. It is a procedure for reducing complexity of correlational data. From the original set of variables under consideration, it selects a smaller set of orthogonal reference axes to span the original data. This technique is also eliminated from consideration. Factor analysis, like principal components, addressess one set of variables and one population.

Cluster analysis is a process of sorting entities into categories according to their overall simularities by comparing vectors of variables. In cluster analysis, very little is known about the category structure.

All that must be known is that there is a collection of observations that are related in some manner. Normally, the operational objective is to discover a category structure which fits the observations. In some situations, it is possible to reduce a very large body of data to a relatively compact description. Cluster analysis is also rejected from consideration. The assumptions that were made before this technique was allowed to be considered were based upon subjective information. The procedures used in cluster analysis to reduce the number of variables is contingent upon seed points, subjectively established, and the subjective interpretation of the overall results of the cluster anlaysis computer programs. The subjective nature of the cluster analysis techniques would compound the subjectiveness already built into the methodology by these assumptions.

Discriminant analysis treats the problem of attempting to differentiate between two or more classes of persons or objects. It attempts to find a linear combination of variables such that the distribution for the classes or groups possess "little" overlap. This technique is accepted for use because it does address the problem of two populations, it does have a form that permits the selection of critical attributes, and it does have a means by which a degree of criticality can be assessed.

The technique for determining critical attributes is stepwise discriminant analysis. It is used to identify a subset of variables which "best" discriminates between populations. It is only necessary for the purposes of this paper to discriminate between two populations, acceptable and unacceptable. Therefore, only the special case involving two populations will be related in this paper.

CHAPTER IV

DETERMINING CRITICAL ATTRIBUTES (METHODOLOGY)

Development

Stepwise discriminant analysis has as its foundation one-way analysis of variance testing means and the discriminant function. These concepts are briefly explained in Appendix A in conjunction with a detailed explanation of stepwise discriminant analysis.

Once these underlying concepts of the multivariate analysis technique are understood, an overall methodology can be developed from all of the previous information. Stepwise discriminant analysis provides a means by which a subset of variables can be identified that "best" discriminates between two populations. These two populations must be defined by individual mean vectors and a common covariance matrix. One sample of observations is given in the form of frequency distributions for the attributes. The attributes to be considered must be designated and the mean and variance for each attribute must be calculated.

Since there is not sufficient information in the original data to formulate acceptable and unacceptable mean vectors for the stepwise discriminant analysis, these vectors will have to be obtained subjectively. There also is not sufficient information from which to formulate the covariance matrix; however, the sample variance of the original data is known. Hence, if the correlations between the variables can be estimated, a covariance matrix can be formulated.

By generating multivariate normal distributions utilizing these mean vectors and the covariance matrix, the stepwise discriminant analysis will enable those attributes which "best" discriminate between the acceptable and unacceptable populations to be identified. The methodology is best presented in the form of steps of a procedure:

Step 1: Examination and preparation of data

Step 2: Determination of the covariance matrix

Step 3: Determination of the mean vectors

Step 4: Generation of the multivariate observations

Step 5: Stepwise discriminant analysis

Step 6: Analysis of results

Explanation of Procedures

Examination and Preparation of Data

The attributes of each test must be examined to determine which level of EEA and which sets of attributes are to be examined. Each test is different from every other test. Even if a certain test is designed to test a system that has been previously evaluated, there will be differences because of the refinements of that first test. In most tests, such as FM 286, the attributes to be analyzed would be broken down into sets of attributes. This would be done because: (1) the number of attributes is so large that the subjective analysis of step 2 would be incomprehensible, (2) the number of attributes is so large that the available computers could not handle the storage required for the preparation of data and the stepwise discriminant analysis, and (3) the attributes were derived from the operational issues in such a manner that natural groupings of attributes would be present.

Having decided the grouping or division of data, each group must be examined to insure that the data for each attribute is available and is in the correct form. The data for each attribute should be represented as a frequency distribution with a range of 1 to 5 that coincides with the rating categories. If an attribute does not have a frequency distribution of observations, then it cannot be evaluated. This procedure does not allow attributes without these data because there is no means to generate a frequency distribution with the information available in the test methodology.

Once the attributes have been organized in the form of distributions, the sample mean and sample variance are calculated. This can be accomplished quite easily by the use of a computer program such as that shown in Appendix B. This program also tests the sets of attributes to insure that the assumption of multivariate normality required for stepwise discriminant analysis is met. In a discussion of multivariate normal distributions, three classes of distributions are of interest. Marginal distributions are the univariate distributions for the individual elements of the vector variable. Conditional distributions are the predicted distributions for particular marginal elements given the known distributions of the remainder of the vector variable. Component distributions are the distributions of any linear functions of the vector variable.

If a vector variable, a vector of attributes, has a multivariate normal distribution, m.n.d., then every one of its marginal distributions is normal. However this is not reversible. If the marginals of the variables of a vector are normally distributed, the vector is not necessarily a m.n.d. (9). If a vector variable has a m.n.d., then every conditional distribution defined on it is normal (9). If a vector variable has a m.n.d., then each component is normally distributed. If every possible linear component of a vector variable is normally distributed, then the vector variable has a m.n.d. (9).

There are no universal goodness of fit tests for multivariate normal distributions (9). Test of multivariate normality have been presented by Malkovich and Afifi (1973), but they are valid for only a small number of variables (33). Although the presence of marginal normality does not insure multivariate normality, this methodology will test for marginal normality. Each marginal of the vector is a special component defined by setting a unit weight for the assigned element and zero weights for all other elements (9). Thus, it is felt that a "better" approximation of multivariate normality can be obtained if the marginals are normally distributed.

In testing the marginals for goodness of fit, it was discovered that the Kolmogorov-Smirnov (K-S) test is more appropriate than the Chi-Square test for the distribution under consideration. The K-S test is considered more appropriate because the power of the test is greater when testing for normality with μ and σ^2 estimated by $\overline{x}$ and s^2 (Afifi & Azen) and because Chi-Square tests conducted on samples of data from test FM 286 showed this procedure infeasible. In the Chi-Square test where the Chi-Square statistic is given by

$$\chi_o^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

the accuracy of the Chi-Square approximation improves as E_i increases. Using five as the minimal acceptable level of E_i, the number of intervals has to be reduced to three. Thus $\chi^2_{\alpha,k-p-1}$ equals zero where k equals the number of intervals and p equals the number of estimated parameters. The hypothesis that the variable conforms to the hypothesized density is rejected if $\chi_o^2 > \chi^2_{\alpha,k-p-1}$. This will always be the case unless O_i equals E_i for each i. This occurrence is highly unlikely.

The K-S test is nonparametric and exact for all sample sizes. In the K-S test, n observations are ordered from smallest to largest. Letting $x_{(i)}$ denote the $i^{\underline{th}}$ smallest observation in the sample, construct the empirical cumulative distribution function $\hat{F}(x)$ defined by

$$\hat{F}(x) = \begin{cases} 0 & , \quad x<x_{(1)} \\ \frac{i}{n} & , \quad x_{(i)} \leq x < x_{(i+1)} \quad i = 1,\ldots,n-1 \\ 1 & , \quad x \geq x_{(n)} \end{cases}$$

The test statistic $D = \frac{\max}{x} \left|\hat{F}(x) - F_o(x)\right|$ tests the null hypothesis, Ho: $F(x) = F_o(x)$. The critical value, D_α, for significance level is established from Table A6 of Fishman (22). Reject the null hypothesis if $D_\alpha < D$. A computer program which facilitates making the K-S test is given in Appendix B.

If the K-S test for normality shows that a frequency distribution for one or more of the attributes does not approximate a normal distribution, then a transformation to induce normality can be utilized. If the transformed data has been tested for the normal fit and the hypothesis has failed to be rejected, the transformed data can be utilized in the remainder of the test. If no transformation to induce normality is found, the remainder of the procedure can be completed with or without this attribute. This point can then be addressed in the final analysis.

Determination of the Covariance Matrix

The determination of a covariance matrix that approximates an actual covariance matrix, unobtainable from actual data, is an essential step in the process of identifying critical attributes. A subjective estimate of covariance can be formulated utilizing the basic relationship of the simple correlation coefficient to the covariance of two variables:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$$

If the correlation between any two variables can be determined, then the individual variances of those variables (obtained in step 1) can be utilized. Thus the problem reduces to that of estimating correlation coefficients for each pair of variables, or in this case, attributes.

The underlying assumption for stepwise discriminant analysis was that the vector of attributes under consideration has a multivariate normal distribution. The multiple correlation coefficient is much more complicated than simple correlation. For the purpose of explanation, let Y represent any of the p+1 attributes and $X_i, i=1,\ldots,p$ represent the remaining p attributes. Also, let the means and variances of the attributes $Y, X_1, \ldots, X_p$, be denoted $\mu_y, \mu_1, \ldots, \mu_p$ and $\sigma_y^2, \sigma_1^2, \ldots, \sigma_p^2$, respectively. Denote the covariance of Y and X_i by σ_{yi}. Correspondingly, the simple correlation coefficients will be defined as given above. These simple correlations do not take into account the presence of more than two attributes. Since the vector of attributes normally contains more than two attributes, the correlation of Y and X_i are actually conditional to the values assumed by the X_j, $j \neq i$.

Let $x_1, \ldots, x_p$ be observations of $X_1, \ldots, X_p$. There exists a conditional distribution of Y given $X_1 = x_1$, $X_2 = x_2, \ldots, X_p = x_p$. In step 1 it was found that if a vector of attributes has a m.n.d., then the conditional distributions are normally distributed. The conditional distribution of Y has mean

$$\mu_{y \cdot x_1, \ldots x_p} = \mu_y + \beta_1(x_1 - \mu_1) + \ldots + \beta_p(x_p - \mu_p)$$

This is called the conditional expectation of Y given $X_1, \ldots, X_p$ (Afifi & Azen). The quantities $\beta_1, \ldots, \beta_p$ are functions of the variances and covariances of the attributes. This conditional distribution has variance

$$\sigma^2 = \sigma_y^2\,(1 - \rho^2_{y \cdot x_1, x_2, \ldots, x_p})$$

where $\rho_{y.x_1,\ldots,x_p}$ is the multiple correlation coefficient of Y and X_1, $\ldots,X_p$ (Afifi & Azen). Transferring $\rho_{y.x_1,\ldots,x_p}$ to the left hand side of the equation by simple algebra results in:

$$\rho^2_{y\cdot x_1,\ldots,x_p} = \frac{\sigma_y^2 - \sigma^2}{\sigma_y^2}$$

Thus, the squared multiple correlation coefficient is equal to the proportion of the variance of attribute Y that is "explained" by the linear relationship with $X_1,\ldots,X_p$. The multiple correlation is the maximum simple correlation between attribute Y and any linear combination of the remaining p attributes. This multiple correlation coefficient is invariant to changes in scale (1).

The discussion of the multiple correlation coefficient illustrates that the relationship between any two variables is highly dependent upon the remaining variables in the vector of attributes. The relationship between Y and X_i is dependent upon the "effect" created by the remaining p-1 variables. With this relationship in mind, let us continue to the actual theory behind finding the correlation between two variables.

In the case of a vector of more than two attributes, multivariate normality was assumed if all of the marginals were normal. Here, if two marginals are normal, the assumption will be that the joint distribution is bivariate normal. A second assumption will be that the values of the remaining attributes in the vector can be set or adjusted to any appropriate levels.

In a bivariate normal distribution, let X_1 and X_2 be distributed normally with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2, respectively. The conditional distribution of X_2 given that $X_1 = x_1$, is univariate normal (Hines & Montgomery) with conditional mean

$$\mu_{2\cdot 1} = \mu_2 + \frac{\sigma_{12}}{\sigma_1 \sigma_2} (x_1 - \mu_1)$$

$$\mu_{2\cdot 1} = \mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x_1 - \mu_1) \tag{1}$$

and conditional variance

$$\sigma^2 = \sigma_2^2 (1 - \rho^2)$$

From equation (1):

$$\frac{\mu_{2\cdot 1} - \mu_2}{\sigma_2} = \rho \frac{(x_1 - \mu_1)}{(\ \sigma_1\ )} \tag{2}$$

The conditional mean, $\mu_{2\cdot 1}$, is the expected value of X_2 given the value of X_1, and μ_1 is the expected value of X_1. The left hand side of equation (2) is the standard normal of conditional X_2 and the expression in the right hand side is a constant times the standard normal of X_1. Thus, the expected value of X_2 given X_1 is ρ times the value of X_1 after normalizing X_2 and X_1.

If μ_1, μ_2, σ_1, σ_2, and ρ are known, then for any given X_1, the value of $\mu_{2\cdot 1}$ can be determined. Letting

$$Z_2 = \frac{(\mu_{2\cdot 1}-\mu_2)}{(\ \sigma_2\ )} \text{ and } Z_1 = \frac{(x_1-\mu_1)}{(\ \sigma_1\ )}, \quad \text{then}$$

$Z_2 = \rho Z_1$. Since Z_1 and Z_2 are N(0,1), if Z_1 is k standard deviations above its mean, then the expected value of Z_2 will be ρ k standard deviations above its unconditional mean. If ρ is known and k varies from the mean, then for each value of k, there is a corresponding value for Z_2. If Z_2 is known, then $\mu_{2\cdot 1}$ can be easily determined. Conversely, if k is specified and Z_2 is estimated, then ρ can be calculated. By letting k vary and estimating corresponding values of Z_2, an estimate of ρ can be obtained by taking a simple average of the ρ's corresponding to the k's. This is by no means an exact process; however, it is a logical approach to estimating the correlation coefficient.

The actual procedure for estimating ρ approximates μ_1 by $\overline{x}_1$, μ_2 by $\overline{x}_2$, σ_1 by s_1, and σ_2 by s_2 obtained in Step 1. As mentioned previously, the relationship between any two variables is contingent upon the values of the remaining variables in the vector. A more exact process would allow the "other" variables to assume all combinations of values over the entire range of the rating categories. A total enumeration, in this case, would be quite infeasible; consequently, the values of X_i calculated in Step 1 will be considered as the levels of the "other" variables.

The actual procedure for the estimation of ρ_{ij} will consider each pair of variables as follows:

1. Consider each of the variables to be normalized, i.e., $Z_i \sim N(0, 1)$

and $Z_j \sim N(0,1)$.

2. Let the scale of measurement range from 1 to 7 corresponding to -3 to +3 standard deviations from the mean, zero.
3. For values of X_i ranging by integers from 1 to 7, excluding 4, estimate the value of X_j on the same scale of 1 to 7 where 1 is considered "best", 4 is considered "borderline", and 7 is considered "worst."
4. Convert the values obtained above from the scale of 1 to 7 to the corresponding standard deviation, -3 to +3.
5. Let the values for X_i equal k. Let the values for X_j equal c. For each value of k, compute a corresponding ρ. This is calculated by $\rho = \frac{c}{k}$.
6. Sum each of the ρ's, and divide by the total number of estimates, 6.
7. Thus, the average of the ρ's is the estimated ρ_{ij} for X_i and X_j.

The estimations should be performed by the individual(s) designated by the agency that is utilizing the overall methodology.

Once all pairwise comparisons are made and all ρ_{ij} are estimated, the esimated covariance matrix is calculated by utilizing the relationship of the simple correlation coefficient to the covariance of two variables:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$$

Letting $\tilde{\rho}_{ij}$ be the estimate of ρ_{ij}, s_i and s_j be estimates of σ_i and σ_j, then s_{ij}, the estimate of σ_{ij} is calculated as follows:

$$s_{ij} = \tilde{\rho}_{ij}\, s_i s_j, \text{ for } i \neq j$$

$$\underset{\sim}{S} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ & \vdots & & \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{bmatrix}$$

where $s_{ii} = s_i$.

This procedure by which the covariance matrix is estimated has limitations because of the assumptions that were made. However, it does reinforce the stand taken by OTEA that subjective evaluations play an extremely valuable role in the design and overall evaluation of operational tests.

<u>Determination of Mean Vectors</u>

The determination of the acceptable and unacceptable mean vectors to be utilized in the stepwise discriminant analysis phase of the methodology to identify critical attributes is vital. The final results hinge upon the acceptable and unacceptable mean values of the individual attributes. For this reason, it is extremely important for those values to be determined by the most knowledgeable and capable individuals involved in the operational tests. These individuals are involved with the development of the test design.

The last step in the development of the Final Test Design is the development of analysis logic. The development of analysis logic is a determination of how the data from the field test will be used to satisfy the test objectives. The analysis methodology includes how the data values which were obtained in the form of observations of measurable data requirements are combined. Data values are combined after considering two types of rules, criteria and weighting.

In order to assess how each aspect of a system's capability performs, a set of criteria for each data requirement at each level is determined. The process of deciding the criteria is difficult, but it can be facilitated by taking into consideration experience, values derived from the test, or comparisons.

Weighting is the importance of each data value expressed as a relationship to the other data values. Weighting takes into account the relative importance of the data. Data values may be given verbal weights such as "essential" or "desirable" or numerical weights.

Those individuals that develop the analysis logic utilizing criteria and weighting are, in effect, determining the basis on which a system and its subsystems are considered acceptable or unacceptable. It is these individuals that should set an acceptable and unacceptable mean level for each data requirement.

A procedure for the determination of the acceptable mean vectors is not tendered here. The exact process by which the criteria and weighting are developed for the data requirements of a particular test are peculiar to that test. Hence, the acceptable and unacceptable mean values for the attributes of each vector are dependent upon the analysis conducted by the test designers.

Generation of Data

Once the covariance matrix and acceptable and unacceptable mean vectors have been determined, it is necessary to generate two sets of multivariate normal data in order to apply stepwise discriminant analysis. In order to generate a multivariate normal distribution, it is necessary to use a fundamental theorem of multivariate statistical analysis. This theorem states that if $\underset{\sim}{Z} = (Z_1,\ldots,Z_p)'$ is $N(\underset{\sim}{0},\underset{\sim}{I})$, then $\underset{\sim}{X}$ with mean $\underset{\sim}{\mu}$ and covariance matrix $\underset{\sim}{\Sigma}$ can be represented as

$$\underset{\sim}{X} = \underset{\sim}{C}\,\underset{\sim}{Z} + \underset{\sim}{\mu}$$

where $\underset{\sim}{C}$ is a unique lower triangular matrix satisfying

$$\underset{\sim}{\Sigma} = \underset{\sim}{C}\,\underset{\sim}{C}'$$

The generation of $\underset{\sim}{X}$ can now be accomplished by: (1) computing $\underset{\sim}{C}$, (2) generating p independent normal variates, and (3) applying $\underset{\sim}{X} = \underset{\sim}{C}\,\underset{\sim}{Z} + \underset{\sim}{\mu}$ (22). Appendix A contians a computer program for the generation of a multivariate normal distribution.

Stepwise Discriminant Analysis

Having generated the necessary data, a stepwise discriminant analysis as discussed in the methodology section is conducted. The F to include, F to exclude, and the significance level α are the means by which the degree of criticality can be somewhat controlled. If F to include is set "too high", the number of variables entering the set of attributes that discriminates between the populations, H, will be severely restricted. If the F to exclude is set "too high", variables will be removed from the set H. If the α level is set "too low", then the

$F_{1-\alpha,\nu_1,\nu_2}$ level for each step will be "too high" and it will affect entry into H.

The F and α values may be adjusted to fit the data. This permits flexibility in the type and quality of the results.

Analysis of Results

This final step in the procedure to identify critical attributes reviews the subjective inputs, the control parameters, and the final results of the stepwise discriminant analysis. These factors are analyzed in an effort to solidify the end results into a productive package that can be practically utilized in future operational tests.

CHAPTER V

DEMONSTRATION OF THE METHODOLOGY

Introduction

This chapter will demonstrate the methodology presented in Chapter IV. The basis for the demonstration will be the data from the Division Command Post Test FM 286. Recall that the test objectives, to evaluate the efficiency of the command post in command and control of division tactical operations and to evaluate the vulnerability of the command post during division tactical operations, were derived from operational issues. These test objectives were further subdivided functionally, as shown by Figure 3, and operationally, as illustrated by Figures 1 and 2. For the purpose of clarity, the demonstration will be restricted to the operations subsystem.

Conduct of the Methodology

Examination and Preparation of Data

The operations subsystem is subdivided into three second level essential elements of analysis. They include:

1. Efficiency in plans, orders, and reports
2. Efficiency in providing information
3. Efficiency in supervising the execution of plans and orders

These second level EEA are related; however, they are still relatively distinct data requirements. At this point, a grouping of the data requirements can be initiated. The fourth level EEA are the measurable data requirements, or attributes. Select the efficiency to provide information, second level EEA, as the example group. In order to restrict the size of the problem for demonstration, select five variables for comparison. The only restriction that is made on the choice of these variables is that there must be data on these attributes in the form of frequency distributions. The five attributes selected and their frequencies are shown in Table 5.

The data on the five attributes are examined. The sample mean, sample variance, sample standard deviation, and maximum Kolmogorov-Smirnov statistics are calculated. If the max K-S statistic is less than a critical level, D_{α}, then the frequency distributions can be accepted as normal. For an α-level of .05, $D_{.05}$ equals $.886/\sqrt{N}$, where N is the total number of observations for an attribute. The critical $D_{.05}$ levels are as follows:

1. Attribute A: .0666
2. Attribute B: .0666
3. Attribute C: .0668
4. Attribute D: .0757
5. Attribute E: .0709

Readily, it is apparent that the original data is rejected as being univariate normal for each attribute, see Table 6. Tables 7 and 8 also show that logarithmic and square root transformations fail to get the data in a form such that the distributions are "acceptable" as normally distributed.

Table 5. Data Distribution

Data Requirement 1.2.2

VARIABLE	RATING CATEGORY					TOTAL
	1	2	3	4	5	
Relevancy of Information	88	73	14	0	2	177
Accuracy of Information	60	95	18	4	0	177
Timeliness of Information	13	49	38	23	5	176
Chg of Com Loc	62	38	22	3	11	137
Organ Concept	24	69	33	10	20	156

Table 6. Data, No Transformation

Variable	Sample Mean	Sample Std. Deviation	Sample Variance	Max K-S Statistic
A	1.6158	0.7303	0.5334	0.5912
B	1.8079	0.7050	0.4970	0.5573
C	1.9432	1.4725	2.1682	0.2090
D	1.9781	1.2094	1.4628	0.4169
E	2.5705	1.2081	1.4595	0.2778

Table 7. Data, Log Transformation

Variable	Sample Mean	Sample Std. Deviation	Sample Variance	Max K-S Statistic
A	0.3910	0.4135	0.1710	0.5256
B	0.5151	0.3995	0.1596	0.4917
C	0.6571	0.5376	0.2890	0.2747
D	0.5283	0.5480	0.3003	0.4153
E	0.8342	0.4782	0.2287	0.2238

Table 8. Data, Square Root Transformation

Variable	Sample Mean	Sample Std. Deviation	Sample Variance	Max K-S Statistic
A	1.2427	0.2682	0.0710	0.4994
B	1.3194	0.2599	0.0676	0.5000
C	1.1664	0.7654	0.5859	0.4922
D	1.3463	0.4084	0.1668	0.4926
E	1.5606	0.3685	0.1358	0.4936

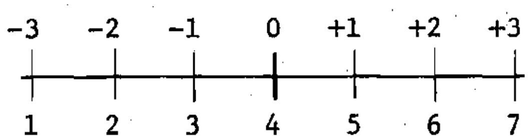
There is a wide disparity between the critical D levels and the maximum K-S statistics. As shown by Figure 5, the distribution "appears" to have a somewhat normal shape. One reason for the disparity is the limited range of observations. A second reason is that the attributes are rated on a continuum from 1 to 5; however, the actual ratings are restricted to integers. With these considerations in mind, the procedure progresses to the determination of the covariance matrix.

Determination of the Covariance Matrix

The first phase in the determination of the covariance matrix process involves "estimating" a correlation coefficient for each pair of attributes.

Step 1: Select a pair of attributes for consideration. Normalize each attribute so that the means of each attribute are equal. Select variables A and C.

Step 2: Let the scale of measurement range from 1 to 7 corresponding to -3 to +3 standard deviations from the mean, zero.



Step 3: Let A vary from 1 to 7 by integers, consider all other attributes to be at their unnormalized means, and estimate C for every value of A.

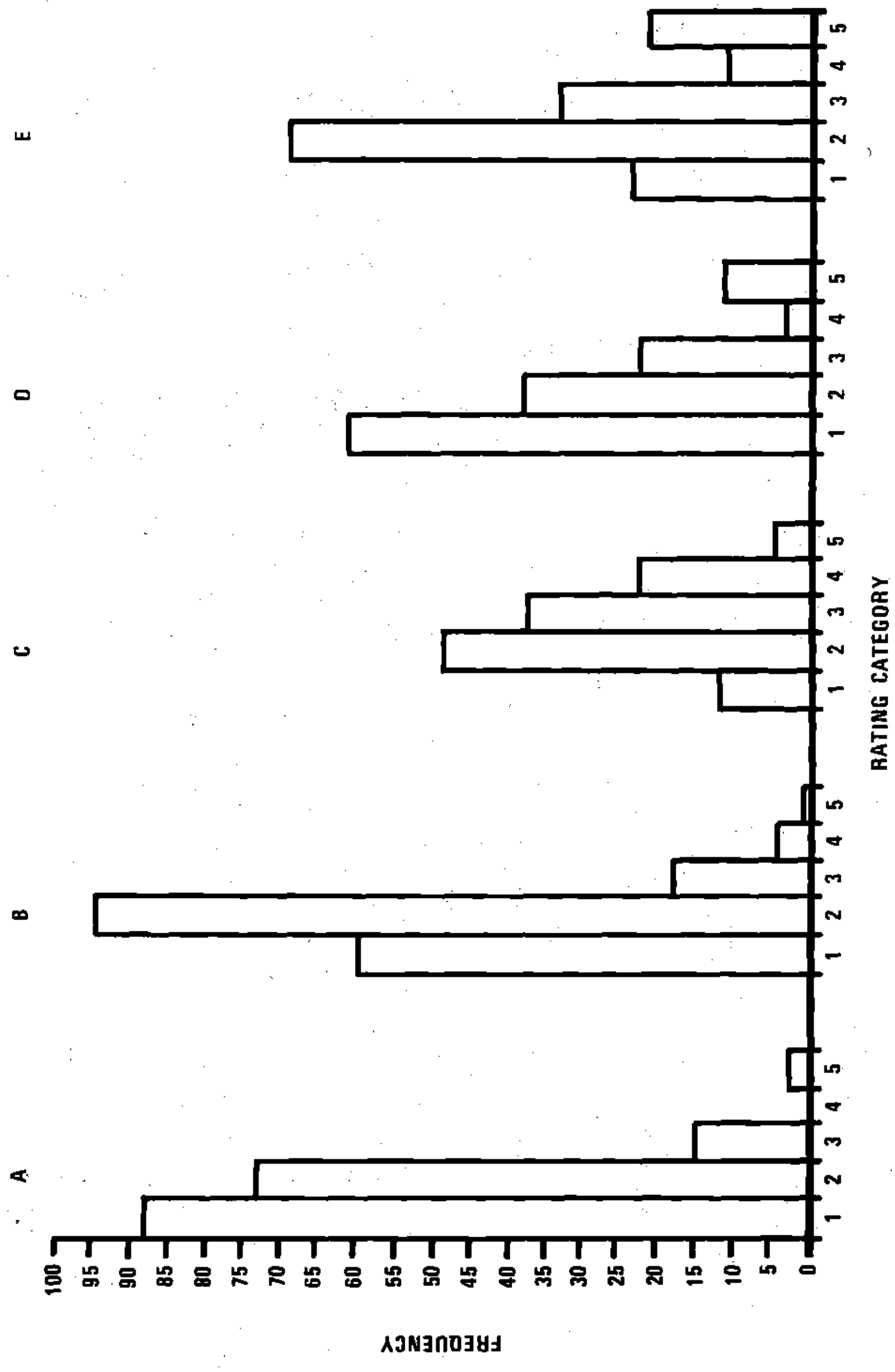


Figure 5. Distribution of Variables

A Value	C Value
1	1
2	1
3	1
-	-
5	2
6	2
7	3

Step 4: Convert the values obtained from Step 3 to standard deviations.

A (k)	C (c)
-3	-3
-2	-3
-1	-3
-	-
+1	-2
+2	-2
+3	-1

Step 5: Let the values for A equal k and the values for C equal c as shown above.

Use $\rho = c/_k$ and calculate the ρ's.

k	c	ρ
-3	-3	1
-2	-3	3/2
-1	-3	3
-	-	
+1	-2	-2
+2	-2	-1
+3	-1	-1/3

Step 6: Sum the ρ's and divide by 6.

Thus $\rho_{AC} = 13/36 = 0.3611$.

Step 7: Finally, use the simple correlation coefficient to solve for S_{AC}.

$$S_{AC} = \rho_{AC} S_A S_C = (0.36)\ (0.7303)\ (1.4725)$$
$$= 0.387$$

Complete the pairwise comparisons of each pair and calculate the "estimated" covariance; see Tables 9 and 10.

Determination of the Mean Vector

The methodology calls for the determination of the acceptable and unacceptable mean vectors by a subjective analysis of the variables involved. A mean vector consisting of 1.5 for each attribute in the acceptable population and a mean vector consisting of 2.5 for each attribute in the unacceptable population were chosen. In this instance, the values were chosen to show what would result from selecting a mean vector on either side of the sample means.

Generation of Data

Two sets of data were generated. One set of data had as its mean, the acceptable mean vector. The second set had as its mean, the unacceptable mean vector. Both sets of data used the covariance matrix in the generation of data. The printout of data is given in Appendix C.

Table 9. Correlation Matrix

VARIABLES	VARIABLES				
	A	B	C	D	E
A	1.00	0.65	0.36	0.58	0.25
B	0.65	1.00	0.47	0.50	0.50
C	0.36	0.47	1.00	0.42	0.65
D	0.58	0.50	0.42	1.00	0.80
E	0.25	0.50	0.65	0.80	1.00

Table 10. Covariance Matrix

VARIABLES	VARIABLES				
	A	B	C	D	E
A	0.5334	0.335	0.387	0.512	0.221
B	0.335	0.497	0.488	0.426	0.426
C	0.387	0.488	2.168	0.748	1.156
D	0.512	0.426	0.748	1.463	1.169
E	0.221	0.426	1.156	1.169	1.460

Stepwise Discriminant Analysis

Biomedical Computer Program 07M utilized the data generated in the previous step. The F to include was set at 0.01 and the F to exclude was set at 0.005. The summary table is as follows:

Table 11. Summary

Step Number	Varible Entered	F Value to Enter	Number of Variables Included	U Statistic
1	2	111.2828	1	0.6402
2	1	14.0945	2	0.5974
3	5	5.8616	3	0.5801
4	4	3.4448	4	0.5700
5	3	1.4660	5	0.5688

The results of each step in the program are given in Appendix C.

Using an α level of 0.05, the summary table is now used to find which attributes should be used in the "best" classification procedure.

F Statistic	F to enter	Variable
$F_{1-\alpha,\nu_1,\nu_2} = F_{.95,1,198} \simeq 3.9$	111.2828	2
$F_{1-\alpha,\nu_1,\nu_2} = F_{.95,1,197} \simeq 3.9$	14.0945	1
$F_{1-\alpha,\nu_1,\nu_2} = F_{.95,1,196} \simeq 3.9$	5.8616	5
$F_{1-\alpha,\nu_1,\nu_2} = F_{.95,1,195} \simeq 3.9$	3.4448	4
$F_{1-\alpha,\nu_1,\nu_2} = F_{.95,1,194} > 3.84$	1.4660	3

Since the F to enter for variable 4 is less than $F_{.95,1,195,}$ then $H = \{2,1,5\}$, or in the notation used in finding the covariance matrix $H = \{B,A,E\}$. Now, since the third step was the last step in which a variable entered, formulate the linear discriminant function from the coefficients and constants at Step 3.

$$a_i = a_{1i} - a_{2i}$$

$$a_1 = a_{11} - a_{21} = 1.91947 - 3.1652 = 1.24615$$

$$a_2 = a_{12} - a_{22} = 1.67710 - 2.83173 = -1.15463$$

$$a_5 = a_{15} - a_{25} = .65847 - 1.11464 = -.45617$$

$$c = c_2 - c_1 = -\ 9.57256 - (-3.87143) = -5.70113$$

Thus, classify $\underset{\sim}{x}$ into W_1, the acceptable population, if

$$Z = -1.24615x_1 - 1.15463x_2 - .4567x_3 \geq -\ 5.70113$$

where the prior probabilities are equal.

Now calculate the estimated Mahalanobis distance, D_q^2, for each step q.

$$D_q^2 = \frac{q(n_1+n_2)(n_1+n_2-2)}{n_1 n_2 (n_1+n_2-q-1)} F$$

where F is the approximation to the U statistic. F is an exact approximation in this case because there are two populations.

$$D_1^2 = \frac{1(200)(198)}{(100)(100)(198)} (0.6402) = 0.012804$$

$$D_2^2 = \frac{2(200)(198)}{(100)(100)(197)} (05.5974) = 0.040203$$

$$D_3^2 = \frac{3(200)(198)}{(100)(100)(196)} (0.5801) = 0.0606122$$

$$D_4^2 = \frac{4(200)(198)}{(100)(100)(195)} (0.5700) = 0.0812307$$

$$D_5^2 = \frac{5(200)(198)}{(100)(100)(194)} (0.5628) = 0.1020618$$

Now test $H_o : \Delta_q^2 = \Delta_p^2$, which is equivalent to testing to see if the last attributes contribute to the discrimination achieved by the attributes in H. Approximate Δ_m^2 by D_m^2, where m=p or q. Testing

$H_o : D_3^2 = D_4^2$, q=3, p=4

$$F = \frac{(n_1+n_2-p-1)}{(p-q)} \quad \frac{n_1 n_2 (D_p^2 - D_q^2)}{(n_1+n_2)(n_1+n_2-2)+n_1 n_2 D_q^2}$$

$$F = 1.005127$$

$$F_{1-\alpha,p-q,n_1+n_2-p-1} = F_{.95,1,195} \approx 3.9$$

$F_{.95,1,195} > F$, hence "fail to reject" the hypothesis.

Testing

$$H_o : D_3^2 = D_5^2$$

$$F = 1.010309$$

$$F_{1-\alpha,2,194} \approx 4.79$$

$F_{.95,2,194} > F$, hence, "fail to reject" the hypothesis.

Analysis of Results

The results of the stepwise discriminant procedure showed that variables B, A, and E "best" discriminated between the acceptable and unacceptable populations. The test of the estimated Mahalanobis distances showed that the discrimination was significant at the .05 level. These findings reinforce the findings of the stepwise process. Although some authors, i.e., Cooley and Lohnes, point out that marginal normality need not be tested when making an assumption of multivariate normality, the results of the K-S test showed that the distribution of the sampling data was not univariate normal for any attribute. If, in fact, the assumption of multivariate normality is violated, then the validity of the results is questionable. If the subjective development of the covariance matrix is not consistent, then the resulting simulated data is not truly representative of the populations.

There are inadequacies in the process, but the approximation or estimated results are still useful for the purpose for which the process was developed. That purpose is to provide an aid by which a test designer could determine which attributes from a group of attributes "best" distinguish between an acceptable and an unacceptable system.

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

Multivariate analysis techniques may be a valuable aid in determining, operationaly, which attributes are more useful in distinguishing between an acceptable and an unacceptable system or subsystem. The use of multivariate analysis, in general, and discriminant analysis, in particular, are adaptable to "real world" operational tests. However, operational tests and force development testing and experimentation are not structurally designed to facilitate the application of multivariate statistical procedures.

Given the present test structure, a viable approach to determining critical attributes is to incorporate a subjectively determined covariance matrix, subjectively determined acceptable and unacceptable mean vectors, and stepwise discriminant analysis in a methodology. Once the covariance matrix has been determined, the selection of the mean attribute vectors can be varied to solicit different results from the stepwise discriminant analysis. It appears that the closer the acceptable and unacceptable vectors bracket the sample mean, the more sensitive the process is to critical attribute selection.

Limitations

This research was conducted under the premise that the basic test design was not going to be changed. Consequently, it is limited by the assumption of marginal normality for the observations on each attribute (and thus multivariate normality). This thesis did not address the feasibility of altering the basic test design in order to insure the validity of the normality assumption.

The data were analyzed with the sole purpose of determining the critical attributes. The applicability of multivariate analysis techniques in the evaluation of the effectiveness of the command and control system was not considered.

The data were concerned with discrete observations over a relatively small range. The multivariate statistical theory is founded on the assumption of a continuous distribution of observations for each attribute.

This research is further limited by the degree to which individual and staff test designers are able to subjectively evaluate measures of effectiveness and to subjectively establish which mean values contribute to a successful and unsuccessful system. Additionally, the definition of the term "critical attribute" precludes the use of other multivariate analysis techniques.

Recommendations

There should be an emphasis put on test design that would permit the use of multivariate analysis techniques to be more applicable. The applicability of nonparametric statistics in determining the correlations among variables should be investigated. The data collection procedure could be designed to permit the use of serial correlation. It could also be modified to enhance the normality of the distribution of observations for each attribute. A more flexible rating scale should be developed that would enhance critical attribute detection. In order to insure the validity of the normality assumption, the rating scheme could be improved by extending the range of the scale or by converting to some other continuous rating scale.

The methodology developed in this thesis should be implemented in future operational tests. This application would enhance the validity of this technique, as well as other multivariate techniques, and assist in the design of forthcoming operational tests.

Further study needs to be done in the area of the sensitivity of stepwise discrimination to the mean vectors. A guideline for the selection of the mean vectors to facilitate the degree to which critical attributes are selected would be highly useful in future applications of this methodology.

APPENDICES

APPENDIX A

ONE-WAY ANOVA

One-way ANOVA is basically a statistical procedure for testing the equality of several means. The underlying theory is founded upon a linear statistical model by which a number of populations are compared on the basis of observations on one random variable. A detailed discussion of the theory is available in Hines and Montgomery (27). In general, the total variation in the data is partitioned into component parts. These component parts, differences between populations and differences within populations, are used to develop a test statistic.

$$SS_T = SS_B + SS_W$$

SS_T is the total sum of squares; SS_B is the between-population sum of squares; and SS_W is the within-population sum of squares. Let a be the number of populations and N be the total number of observations on a random variable, X. Let the null hypothesis be that the means of the populations are equal. It is assumed that X is normally distributed; thus, SS_T/σ^2 is a Chi-Square distribution with N-1 degrees of freedom, df. Hence, SS_W/σ^2 is χ^2(N-a) and, if the means are equal, SS_B/σ^2 is χ^2 (a-1). Under the null hypothesis, H_o,

$$\frac{SS_T}{\sigma^2} = \frac{SS_B}{\sigma^2} + \frac{SS_W}{\sigma^2}$$

or

$$\chi^2(N-1) = \chi^2(a-1) + \chi^2(N-a)$$

where SS_B/σ^2 and SS_W/σ^2 are independent χ^2 random variables. It is also known that $F_o = \frac{SS_B/(a-1)}{SS_W/N-a)}$ is $F(a-1,N-a)$. Call $SS_B/(a-1)$ and $SS_W/(N-a)$ mean squares, MS_B and MS_W respectively. It can be shown that F_o is an appropriate test statistic by taking the expected value of MS_B and MS_W. If the value of F_o is too large, then the null hypothesis that the means are equal should be rejected. It will be seen in the stepwise discriminant analysis procedure that it is highly desirable to reject the null hypothesis and to have the highest value of F_o that is possible.

DISCRIMINANT FUNCTION

The standard classification procedure for p continuous variables assumes that a vector of observations comes from one of two multivariate normal populations. Let the vector of observations be represented by $\underset{\sim}{x} = (x_1, \ldots, x_p)'$ and assume that one population, W_1 is $N(\underset{\sim}{\mu}_1^{p\times1}, \underset{\sim}{\Sigma}^{p\times p})$ and the second population, W_2, is $N(\underset{\sim}{\mu}_2^{p\times1}, \underset{\sim}{\Sigma}^{p\times p})$.

If $\underset{\sim}{\mu}_1$, $\underset{\sim}{\mu}_2$, and $\underset{\sim}{\Sigma}$ are assumed to be known, then it seems reasonable, intuitively, that a linear combination of the observations can be found by which that vector of observations can be classified into W_1, or W_2. The linear combination of observations

$$Z_1 = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_p x_p \tag{1}$$

is called a discriminant function. Classify $\underset{\sim}{x}$ into W_1 if Z is greater than or equal to some constant C or classify $\underset{\sim}{x}$ into W_2 if Z is less than that constant. Thus, if the α_i and C can be determined such that the probability of making an incorrect classification is minimized, then the problem is solved.

Suppose $\underset{\sim}{x}$ is from W_1. In this case, Z is $N(\zeta_1, \sigma_Z^2)$ where

$$\zeta_1 = \sum_{j=1}^{p} \alpha_j \mu_{1j}, \text{ and}$$

$$\sigma_Z^2 = \sum_{i=1}^{p} \sum_{j=1}^{p} \alpha_i \sigma_{ij} \alpha_j .$$

If $\underset{\sim}{x}$ is from W_2, then Z is $N(\zeta_2, \sigma_Z^2)$ where

$$\zeta_2 = \sum_{j=1}^{p} \alpha_j \mu_{2j}, \text{ and}$$

σ_Z^2 is given above. In order to maximize the distance between the two populations, the α_i should be chosen so that the means of the two populations are as far apart as possible. Thus the Mahalanobis distance

$$\Delta^2 = \frac{(\zeta_1 - \zeta_2)^2}{\sigma_Z^2} \tag{2}$$

can be utilized. It can be shown that the α_i coefficients which maximize Δ^2 are the solutions to the set of linear equations

$$\alpha_1 \sigma_{11} + \alpha_2 \sigma_{12} + \ldots + \alpha_p \sigma_{1p} = \mu_{11} - \mu_{21}$$

$$\alpha_1\sigma_{21} + \alpha_2\sigma_{22} + \ldots + \alpha_p\sigma_{2p} = \mu_{12} - \mu_{22}$$
$$\vdots$$
$$\alpha_1\sigma_{p1} + \alpha_2\sigma_{p2} + \ldots + \alpha_p\sigma_{pp} = \mu_{1p} - \mu_{2p} \tag{3}$$

The discriminant score Z for a vector of observations can be found by using the α_i obtained from the solution of (3) in equation (1).

Intuitively, the constant C would be that point between ζ_1 and ζ_2 that minimized the probability of classifying $\underset{\sim}{x}$ into W_1, or W_2 incorrectly. Since the variance for both populations are equal, then it seems obvious that C should be the midpoint between ζ_1 and ζ_2 (see Figure 6).

Thus the procedure is to classify $\underset{\sim}{x}$ into W_1 if $Z \geq C$ or

$$\sum_{j=1}^{p} \alpha_j x_j \geq \frac{\zeta_1 + \zeta_2}{2}$$

and to classify $\underset{\sim}{x}$ onto W_2 if $Z<C$ or

$$\sum_{j=1}^{p} \alpha_j x_j < \frac{\zeta_1 + \zeta_2}{2}.$$

It can be shown (1) that this intuitive approach is correct if the a priori probability that a vector comes from W_1 is equal to the a priori probability that it comes from W_2 and if the costs of misclassification are equal. Otherwise, from the generalized Bayes classification procedure presented by Afifi and Azen classify $\underset{\sim}{x}$ into W_1 if

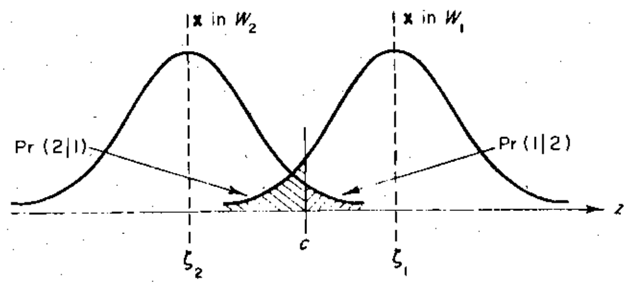


Figure 6. Distribution of Z.

$$Z \geq \frac{\zeta_1+\zeta_2}{2} + \ln \frac{q_2}{q_1} \frac{C(1|2)}{C(2|1)} \tag{4}$$

and classify $\underset{\sim}{x}$ into W_2 if

$$Z < \frac{\zeta_1+\zeta_2}{2} + \ln \frac{q_2\ C(1|2)}{q_1\ C(2|1)} \tag{5}$$

where q_i are the a priori probabilities for being classified into population W_i, i=1,2 and C(2|1) and C(1|2) are the costs of misclassification.

It was assumed initially that the parameters of the population distributions were known. If $\underset{\sim}{\mu}_1, \underset{\sim}{\mu}_2$, and $\underset{\sim}{\Sigma}$ are unknown and if $x_{i1},\ldots,x_{in_i}$, i=1,2 are independent random samples from W_1 and W_2, then $\underset{\sim}{\mu}_1, \underset{\sim}{\mu}_2$, and $\underset{\sim}{\Sigma}$ can be estimated by $\overline{\underset{\sim}{x}}_1, \overline{\underset{\sim}{x}}_2$, and $\underset{\sim}{S}$, respectively, where $\overline{\underset{\sim}{x}}_i$, i=1,2 are sample means and $\underset{\sim}{S}$ is the pooled sample covariance matrix. These consistent estimators are applied to the generalized Bayes classification procedure which becomes an estimated generalized Bayes classification procedure. Using $\overline{x}_{ij}$, i=1,2, j=1,...,p and s_{jm},m=1,...,p in equation (3), solve for estimates of α_i denoted by a_i. Use the a_i to calculate the estimated discriminant score Z_{ik} for each observation vector $\underset{\sim}{x}_{ik}$, k=1,...,n_i. Estimate ζ_i by

$$\overline{Z}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} Z_{ik}$$

and estimate σ_Z^2 by

$$s_Z^2 = \sum_{j=1}^{p} \sum_{m=1}^{p} a_j\ s_{jm}\ a_m.$$

Thus the estimated generalized procedure modifies equations (4) and (5) so that $\underset{\sim}{x}$ is classified into W_1 if

$$Z = \sum_{i=1}^{p} a_i x_i \geq \frac{\overline{Z}_1+\overline{Z}_2}{2} + \ln \frac{q_2\ C(1|2)}{q_1\ C(2|1)} \qquad (6)$$

and $\underset{\sim}{x}$ is classified into W_2 if

$$Z < \frac{\overline{Z}_1+\overline{Z}_2}{2} + \ln \frac{q_2\ C(1|2)}{q_1\ C(2|1)}$$

The estimate of Δ^2 is the sample Mahalanobis distance,

$$D^2 = \frac{(\overline{Z}_1-\overline{Z}_2)^2}{s_Z^2}$$

Under the assumptions that the several original variables have a multivariate normal distribution within the populations from which the samples were drawn and that the covariance matrices for the two populations are equal, an F statistic derived from D^2 can be used to test H_o: $\Delta^2=0$ or equivalently H_o: $\underset{\sim}{\mu}_1=\underset{\sim}{\mu}_2$.

$$F = \frac{n_1 n_2(n_1+n_2-p-1)}{p(n_1+n_2)(n_1+n_2-2)}\ D^2$$

F is compared with $F_{1-\alpha,p,n_1+n_2-p-1}$. Thus, this F statistic is the same as that F which utilizes the two-sample Hotelling T^2 statistic for testing equality of mean vectors.

$$F = \frac{n_1+n_2-p-1}{(n_1+n_2-2)p} \quad T^2$$

STEPWISE DISCRIMINANT ANALYSIS

Stepwise discriminant analysis is a multivariate analysis technique by which a subset of variables that "best" discriminate between k populations can be identified from the whole set of variables. This thesis is only concerned with two populations. The case of k=2 will be discussed here. The procedures for the general case, where k can represent multiple populations, is presented in Afifi and Azen (1). The terminology used by Afifi and Azen is adapted here.

Consider two multivariate normal populations, W_1 and W_2, with parameters discussed above where a p dimensional observation vector was classified into one of two multivariate normal populations. The general logic behind the stepwise discriminant procedure is to first identify the variable for which the mean values in the two populations are "most different." This variable will be entered into a separate set. Thereafter, on successive steps, the conditional distribution of each variable not entered, given the variables entered, will be considered. Of the variables not entered, the variable for each mean value of the conditional distributions in the two populations which are "most different" will be identified. This variable will be added to the separate set. The stepwise process is stopped when no additional variables significantly contribute to the discrimination between the two populations. The measure by which the "most different" variables are selected is a one-way analysis of variance (ANOVA) F statistic.

A one-way ANOVA F statistic that is calculated on those variables not entered into the separate set is called an F to enter. A one-way ANOVA F statistic calculated for those variables that are in the separate set is called an F to remove. There are two fixed F statistics that are used as control parameters in the procedure. These are F to include and F to exclude. These are minimal acceptable values for F to enter and F to remove.

A detailed program for the stepwise procedure is as follows:

Step 0: Let $x_{11}^{px1},\ldots,x_{1n_1}^{px1}$, and $x_{21}^{px1},\ldots,x_{2n_2}^{px1}$ be random samples from W_1 and W_2, respectively. The F to enter along with its degree of freedom (df) is computed for each $X_j, j=1,\ldots,p$. This F is a one-way ANOVA F statistic for testing H_o: $\mu_{1j}=\mu_{2j}$, for $j=1,\ldots,p$. If all F to enter are less than F to include, a prescribed inclusion level, the process is terminated. If this occurs, the conclusion is that no variable significantly discriminates between the two populations.

Step 1: The variable X_{j_1} having the largest F to enter is chosen as the first variable to enter the separate set, H. The estimated linear discriminant coefficient and constant are calculated for both population W_1 and W_2. The classification table, U statistic, and an F approximation to U are calculated. The F to remove for X_{j_1}, which is equal to the F to enter, and its df is calculated. The F to enter and its df for each variable not entered are calculated. These test the hypothesis H_o: $\mu_{1j.j_1}=\mu_{2j.j_1}$ where $\mu_{ij.j_1}$ is the mean of the conditional distribution in W_i of X_j given X_{j_1}, $i=1,2$, $j=1,\ldots,p$, $j\neq j_1$. If all the F to enter are less than F to include, then the last step, Step S, is executed. Otherwise, Step 2 is executed.

Step 2: The variable X_{j_2} is chosen from those F to enter computed in step 1 that has the maximum F. Thus, now the separate set contains X_{j_1} and X_{j_2}, $H = \{X_{j_1}, X_{j_2}\}$. The two estimated linear discriminant coefficients and constant are calculated for each population W_1 and W_2. The classified table, U statistic, and F approximation to U are calculated. The F to remove and their df are calculated for X_{j_1} and X_{j_2}. These test H_o: $\mu_{1j_1.j_2}=\mu_{2j_1.j_2}$ and H_o: $\mu_{1j_2.j_1}=\mu_{2j_2.j_1}$, where $\mu_{ij.j_1j_2}$ is the mean of the conditional distribution in W_i of X_j given X_{j_1} and X_{j_2}, $i=1,2$, $j=1,\ldots,p$, $j\neq j_1$ or j_2.

Step 3: (a) Letting L denote the set of k variables which have been entered (replacing H). If any of the F to remove for the variables in L are less than F to exclude, a prescribed deletion level, delete from L the variable with the smallest F to remove and execute (b) with k-1 replacing k. If all the F to enter for the variables not in L are less than F to include, execute Step S. Otherwise, choose the variable with the largest F to enter, and add it to L so that k+1 replaces k.

(b) The k estimated linear coefficients and constant are calculated for W_1 and W_2. The classification table, U statistic, and F approximation to U are calculated. Calculate the F to remove and its df for each variable in L. These test $H_o:\mu_{1s.(k-1)} = \mu_{2s.(k-1)}$ for each X in L given the remaining k-1 variables in L. The notation $\mu_{is.(k-1)}$ is the mean of the conditional distribution in W_i of X_s given all the other variables in L except X_s. The F to enter and their df are calculated for those variables not in L. These test H_o: $\mu_{1j.k}=\mu_{2j.k}$, where $\mu_{ij.k}$ is the mean of the conditional distribution in W_i of X_j given all the variables in L where $i=1,2$, $j=1,\ldots,p$, X_j not in L.

Step 4,5,...: Repeat Step 3 recursively until all the variables have been entered and no F to remove is less than the F to exclude or when the F to enter is less than the F to include for all variables not in L. Then execute Step S.

Step S: The posterior probability of belonging to W_1 and W_2 is calculated for $\underset{\sim}{x}_{1k}$ and $\underset{\sim}{x}_{2k}$ where k designates the variables entered. These probabilities are used to classify each set of data into one of the two populations and a classification table is prepared. A summary table is prepared. In this table, the step number, the variable entered or removed, the F to enter or F to remove, the U statistic, and the F approximation to U is given for each step.

This completes the stepwise discriminant analysis procedure. The summary table is then used to determine which variables "best" discriminate between the two populations. Let α be a prescribed significance level. If the F to enter for variable X_{j_1} in Step 1 is less than $F_{1-\alpha,\nu_1,\nu_2}$, then it is not possible to significantly discriminate between the two populations. Otherwise, H = $\{X_{j_1}\}$. In Step 2, X_{j_2} was entered. If its F to enter is less than $F_{1-\alpha,\nu_1,\nu_2,}$ then H = $\{X_{j_1}\}$. Otherwise, H = $\{X_{j_1},X_{j_2}\}$. For each step thereafter, if a variable is removed, delete it from H and go on to the next step. If a variable is entered, compare its F to enter with $F_{1-\alpha,\nu_1,\nu_2}$. If its F to enter is less than $F_{1-\alpha,\nu_1,\nu_2}$, stop amending H. Otherwise, augment H and go on to the next step.

Once all variables that are eligible for entry have entered H, the linear discriminant function, d, can be obtained by taking the difference of the two estimated discriminant scores.

$$a_i = a_{1i} - a_{2i}$$

$$C = \frac{\overline{Z}_1+\overline{Z}_2}{2} = C_2 - C_1$$

$$d = \sum_{x_i \varepsilon H} a_i x_i$$

Thus, an observation vector containing those variables in H can be classified into population W_1 if $d \geq C$. Let q designate the variables in H. Since there are only two populations, the F approximation to U is exact. The estimated Mahalanobis distance based on q variables $D_q^{\,2}$ may be obtained from the F approximation to U.

$$D_q^{\,2} = \frac{q(n_1+n_2)(n_1+n_2-2)}{n_1 n_2(n_1+n_2-q-1)} F$$

for q=1,...,p and the sample sizes n_1 and n_2 from W_1 and W_2, respectively. Suppose the variables $X_1,\ldots,X_q$ are in H. To test that the remaining

$X_{q+1}, \ldots, X_p$ do not contribute to the discrimination achieved by the variables in H, test H_o: $\Delta_q^2 = \Delta_p^2$. This tests the difference between the population Mahalanobis distance based on the q variables in H and all p variables. Let D_q^2 and D_p^2 be the sample estimates of Δ_q^2 and Δ_p^2. Then,

$$F = \frac{n_1+n_2-p-1}{p-q} \frac{n_1 n_2 (D_p^2 - D_q^2)}{(n_1+n_2)(n_1+n_2-2) + n_1 n_2 D_q^2}$$

Under H_o, this F has a F_{p-q, n_1+n_2-p-1} distribution. If $F > F_{1-\alpha, p-q, n_1+n_2-p-1}$, reject H_o. If H_o is rejected, this would mean that one or more of the p-q variables not in H, contribute to the discrimination between W_1 and W_2 for a given significance level.

APPENDIX B

```
WILLIAMS-G*KSTEST.WILL
              DIMENSION FREQ(10,10),N(10), XVAL(10), XHAT(10), XVAR(10)
              DIMENSION SVAR(10), SHAT(10), STD(10), X(10,200), F(10,200)
              DIMENSION NN(200), Y(10), FZ(10), D(10), DMAX(10)
              DIMENSION Z(10)
              INTEGER FREQ
              READ(5,55) LC
              READ(5,15) M
           15 FORMAT()
              READ(5,25) NRC

           25 FORMAT()
              DO 40 I=1,M
              READ(5,35) (FREQ(I,J), J=1,NRC)
           35 FORMAT()
           40 CONTINUE
              READ(5,55) (N(I), I=1,M)
           55 FORMAT()
        C*****COMPUTE SAMPLE MEAN AND SAMPLE VARIANCE
              IF(LC.EQ.0) GO TO 57
           56 IF(LC.EQ.1) GO TO 60
              IF(LC.EQ.2) GO TO 63
           57 XVAL(1)=0
              DO 59 I=1,M
                  DO 58 J=1,NRC
                  XVAL(1) = XVAL(I) + (FREQ(I,J)*J/1.0)
           58 CONTINUE
              XHAT(I)=XVAL(I)/(N(I)*1.0)
           59 CONTINUE
              GO TO 66
           60 DO 62 I=1,M
                  DO 61 J=1,NRC
                  XVAL(1) = XVAL(I) + (FREQ(I,J)*LOG(J*1.0))
           61 CONTINUE
              XHAT(I)=XVAL(I)/(N(I)*1.0)
           62 CONTINUE
              GO TO 66
           63 DO 65 I=1,M
                  DO 64 J=1,NRC
                  XVAL(1) = XVAL(I) + (FREQ(I,J)*SQRT(J/1.0))
           64 CONTINUE
              XHAT(I)=XVAL(I)/(N(I)*1.0)
              XHAT(I) = XVAL(I)/(N(I)*1.0)
           65 CONTINUE
           66 XVAR(1)=0
               IF(LC.EQ.1) GO TO 69
               IF(LC.EQ.2) GO TO 72
              DO 68 I=1,M
                  DO 67 J=1,NRC
                  XVAR(1) = XVAR(I) + FREQ(1,J)*((J/1.0)**2)
           67 CONTINUE
              SVAR(I)=(XVAR(I)-(XVAL(I)**2)/(N(I)*1.0))/(N(I)-1)
              SHAT(I)=SQRT(SVAR(I))
           68 CONTINUE
              GO TO 75
           69 DO 71 I=1,M
                  DO 70 J=1,NRC
```

```
               XVAR(I) = XVAR(I) + FREQ(I,J)*(LOG(J)**2)
        70 CONTINUE
           SVAR(I)= (XVAR,I)-(XVAL(I)**2)/(N(I)*1.0))/(N(I)-1)
           SHAT(I)=SQRT(SVAR(I))
        71 CONTINUE
           GO TO 75
        72 DO 74 I=1,M
               DO 73 J=1,NRC
               XVAR(I) = XVAR(I) + FREQ(I,J)*((SQRT(J))**2)
        73 CONTINUE
           SVAR(I)= (XVAR,I) - (XVAL(I)**2)/(N(I)*1.0))/(N(I)-1)
           SHAT(I)=SQRT(SVAR(I))
        74 CONTINUE
C*****STANDARDIZE EACH OBSERVATION AND PUT IT IN AN ARRAY
        75 IF(LC.EQ.1) GO TO 77
           IF(LC.EQ.2) GO TO 79
           DO 76 J=1,NRC
           STD(J)=(J*1.0-XHAT(I))/SHAT(I)
        76 CONTINUE
           GO TO 81
        77 DO 78 J=1,NRC
           STD(J)=((LOG(J*1.0))-XHAT(I))/SHAT(I)
        78 CONTINUE
           GO TO 81
        79 DO 80 J=1,NRC
            STD(J)=((SQRT,J*1.0))-XHAT(I))/SHAT(I)
        80 CONTINUE
        81 DO 95 I=1,M
           L=1
               DO 90 J=1,NRC
           Y(I)= Y(I) + FREQ(I,J)
           MM = 0
           MM=Y(I)
               DO 85 K=L,MM
               X(I,K) = STD(J)
        85               CONTINUE
               L=Y(I) + 1
        90           CONTINUE
               NN(I) = Y(I)
        95 CONTINUE
C**** TEST THE M VARIABLES FOR NORMALITY
           DO 105 I=1,M
               DO 100 K=1,NN(I)
               F(I,K) = (K*1.0) / (NN(I)*1.0)
       100     CONTINUE
       105 CONTINUE
           DO 120 I =1,M
           DMAX(I) = 0
               DO 115 J=1,NN(I)
                FZ(I) = RNORM(X(I,J))
               D(I) = ABS(F(I,J)-FZ(I))
               IF(D(I).LE.DMAX(I)) GO TO 115
               DMAX(I) = D,I)
       115 CONTINUE
       120 CONTINUE
C**** PRINT THE OUTPUT
           DO 200 I = 1,M
```

```
      WRITE(6,130) I
  130 FORMAT(/,11X,15H  VARIABLE  X (,I2,1X,1H))
      WRITE(6,135) XHAT(I)
  135 FORMAT(/,11X,17H  SAMPLE MEAN  = ,F10.4)
      WRITE(6,140) SVAR(I)
  140 FORMAT(/,11X,21H  SAMPLE VARIANCE  = ,F10.4)
      WRITE(6,145) SHAT(I)
  145 FORMAT(/,11X,31H  SAMPLE STANDARD DEVIATION  = ,F10.4)
      WRITE(6,150) DMAX(I)
  150 FORMAT(/,11X,23H  MAX K-S STATISTIC  = ,F10.4,/)
  200 CONTINUE
      DO 210 I=1,M
      Z(I)=.886/SQRT(N(I))
  210 CONTINUE
      LC1=0
      DO 220 I=1,M
      IF(DMAX(I).LT.Z(I))  GO TO 220
      LC1=LC1+1
  220 CONTINUE
       DO 240 I=1,M
           XVAL(I)=0
           XHAT(I)=0
           XVAR(I)=0
            SVAR(I)=0
         SHAT(I)=0
         STD(I)=0
         Y(I)=0
         FZ(I)=0
         D(I)=0
         DMAX(I)=0
          Z(I)=0
  240 CONTINUE
       DO 260 I=1,10
          DO 250 J=1,200
          X(I,J)=0
          F(I,J)=0
          NN(J)=0
  250     CONTINUE
  260   CONTINUE
  999 STOP
      END
```

```
WILLIAMS-G*MULNOR.MAIN
            DIMENSION CT(20,20),CK(20,20)
            COMMON /ONE/U(20)
            COMMON /TWO/SIGMA(20,20)
            COMMON /THREE/CMAT(20,20)
            COMMON /FOUR/ZVEC(20),XVEC(20),BUF(20)
            COMMON /FIVE/NL
            EXTERNAL UNIF,RNORM1
      41    FORMAT(//,5X,'**  COVARIANCE MATRIX  **')
      42    FORMAT(//,5X,'** C  MATRIX  **')
      43    FORMAT(//,5X,'**  MEAN VECTOR  **')
      44    FORMAT(//,5X,'**  GENERATED OBSERVATIONS  **')
      45    FORMAT(//,5X,'**  CHECK MATRIX  **')
      53    FORMAT(8F10.4)
      54    FORMAT(/,2X,8(1X,F8.4))
      55    FORMAT()
            WRITE(6,109)
     109    FORMAT(/,2X,'ENTER THE NR OF STARTUP RUNS FOR UNIF')
            READ(5,55) NSR
            DO 100 IV=1,NSR
            QQ=UNIF(A)
     100    CONTINUE
              WRITE(6,101)
     101    FORMAT(/,2X,'ENTER DIMENSION OF RESPONSE ')
            READ(5,55,END=999) NL
            WRITE(6,103)
     103    FORMAT(/,2X,'ENTER THE SAMPLE SIZE')
            READ(5,55,END=999) NK
            WRITE(6,105)
     105    FORMAT(/,2X,'ENTER THE MEAN VECTOR')
            READ(5,55,END=999)(U(I),I=1,NL)
            WRITE(6,107)
     107    FORMAT(/,2X,'ENTER THE COVARIANCE MATRIX')
            READ(5,55,END=999)((SIGMA(I,J),J=1,NL),I=1,NL)
            CALL CMAT1
            CALL MXTRN(CMAT,CT,NL,NL,20,20)
            CALL MXMLT(CMAT,CT,CK,NL,NL,NL,20,20)
            WRITE(6,43)
            WRITE(6,54)(U(I),I=1,NL)
            WRITE(6,41)
            DO 500 I=1,NL
            WRITE(6,54)(SIGMA(I,J),J=1,NL)
     500    CONTINUE
            WRITE(6,42)
            DO 520 I=1,NL
            WRITE(6,54)(CMAT(I,J),J=1,NL)
     520    CONTINUE
            WRITE(6,45)
            DO 550 I=1,NL
            WRITE(6,54)(CK,I,J),J=1,NL)
     550    CONTINUE
            WRITE(6,44)
            DO 630 K=1,NK
            CALL XVEC1(RNORM1,UNIF)
            WRITE(8,53)(XVEC(L),L=1,NL)
     630    CONTINUE
     999    CONTINUE
            END
```

```
WILLIAMS-G*MULNOR.CMAT
              SUBROUTINE CMAT1
              COMMON /FIVE/N
              COMMON /THREE/CMAT(20,20)
              COMMON /TWO/SIGMA(20,20)
              DO 110 J=1,N
              IF(J.GE.2) GO TO 91
              DO 81 I=1,N
              CMAT(I,1)=SIGMA(I,1)/SQRT(SIGMA(1,1))
       81     CONTINUE
              GO TO 110
       91     DO 105 I=1,N
              IF(J.GE.I+1) GO TO 104
              IF(J.NE.1) GO TO 95
              SUB1=0.0
              L=I-1
              DO 93 K=1,L
              SUB1=SUB1+CMAT(I,K)**2
       93     CONTINUE
              CMAT(I,J)=SQRT(SIGMA(I,J)-SUB1)
              GO TO 105
       95     SUB2=0.0
              L=J-1
              DO 97 K=1,L
              SUB2=SUB2+CMAT(I,K)*CMAT(J,K)
       97     CONTINUE
              CMAT(I,J)=(SIGMA(I,J)-SUB2)/CMAT(J,J)
              GO TO 105
      104     CMAT(I,J)=0.0
      105     CONTINUE
      110     CONTINUE
              RETURN
              END
```

```
WILLIAMS-G*MULNOR.RNORM
              FUNCTION RNORM1(UNIF,RNORM2,U,SIG2)
              TPI=6.2831852
              A=UNIF(X)
              B=UNIF(X)
              RNORM1=U+SQRT(-2.0*SIG2*ALOG(A))*COS(TPI*B)
              RNORM2=U+SQRT(-2.0*SIG2*ALOG(A))*SIN(TPI*B)
              RETURN
              END
```

```
WILLIAMS-G*MULNOR.XVEC
                SUBROUTINE XVEC1(RNORM1,UNIF)
                COMMON /ONE/U(20)
                COMMON /THREE/CMAT(20,20)
                COMMON /FOUR/ZVEC(20),XVEC(20),BUF(20)
                COMMON /FIVE/N
                DO 27 I=1,N,2
                ZVEC(I)=RNORM1,UNIF,RNORM2,0.0,1.0)
                II=I+1
                ZVEC(II)=RNORM2
        27      CONTINUE
                DO 121 I=1,N
                SUM=0.0
                DO 111 J=1,N
                SUM=SUM+CMAT(I,J)*ZVEC(J)
        111     CONTINUE
                BUF(I)=SUM
        121     CONTINUE
                DO 131 K=1,N
                XVEC(K)=BUF(K)+U(K)
        131     CONTINUE
                RETURN
                END
```

```
WILLIAMS-G*MULNOR.UNIF
                FUNCTION UNIF(A)
                DATA IY/96581/
                IY=IY*3125
                IF(IY) 5,6,6
        5       IY=IY+1+343597,8367
        6       YFL=IY
                UNIF=YFL*2.0**,-35)
                RETURN
                END
```

APPENDIX C

BMD07M - STEPWISE DISCRIMINANT ANALYSIS - REVISED MAY 17, 1971
HEALTH SCIENCES COMPUTING FACILITY, UCLA

PROBLEM CODE EXAMP

NUMBER OF VARIABLES 5

NUMBER OF GROUPS 2

NUMBER OF CASES IN EACH GROUP 100 100

PRIOR PROBABILITIES .5000 .5000

VARIABLE FORMAT (5F10.4)

DATA INPUT FROM CARDS

MEANS (THE LAST COLUMN CONTAINS THE GRAND MEANS OVER THE GROUPS USED IN THE ANALYSIS)

VARIABLE	GROUP ACC	UNACC	
1	1.49410	2.48618	1.99014
2	1.50239	2.51641	2.00940
3	1.55932	2.62139	2.09036
4	1.51816	2.50174	2.00995
5	1.47164	2.47858	1.97511

STANDARD DEVIATIONS

VARIABLE	GROUP ACC	UNACC
1	.70121	.69884
2	.67733	.68206
3	1.32037	1.30637
4	1.09468	1.08999
5	1.06626	1.06729

WITHIN GROUPS COVARIANCE MATRIX

VARIABLE	VARIABLES 1	2	3	4	5
1	.49004				
2	.28067	.46199			
3	.24448	.30852	1.72499		
4	.37180	.26824	.29374	1.19320	
5	.12571	.28681	.76773	.88524	1.13800

WITHIN GROUPS CORRELATION MATRIX

VARIABLE	VARIABLES 1	2	3	4	5

1	1.00000				
2	.58988	1.00000			
3	.26591	.34560	1.00000		
4	.48623	.36128	.20475	1.00000	
5	.16834	.39555	.54796	.75968	1.00000

```
SUBPROBLEM                          1
F-LEVEL FOR INCLUSION        .0100
F-LEVEL FOR DELETION         .0050
TOLERANCE LEVEL              .0001
CONTROL VALUES            IIIII

*****************************************************************************

STEP NUMBER              0
VARIABLE ENTERED

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM      1   198

  1 100.4229            2 111.2828            3   32.6957            4   40.5390
  5  44.5483
*****************************************************************************

STEP NUMBER              1
VARIABLE ENTERED         2

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM      1   198

  2 111.2828

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM      1   197

  1   14.0945           3    3.1063           4    4.7855            5    4.7261

U-STATISTIC                .64019     DEGREES OF FREEDOM     1     1   198
APPROXIMATE F           111.28283     DEGREES OF FREEDOM     1   198.00

F MATRIX - DEGREES OF FREEDOM      1   198

              GROUP
               ACC
GROUP
  UNACC    111.28282

            FUNCTION
               ACC            UNACC
VARIABLE
  2         3.25199          5.44687

CONSTANT
           -3.13603         -7.54644

             NUMBER OF CASES CLASSIFIED INTO GROUP -
               ACC    UNACC
GROUP
    ACC        76       24
  UNACC        27       73

*****************************************************************************

STEP NUMBER              2
VARIABLE ENTERED         1
```

```
VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1   197

  1  14.0945          2  21.7764

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1   196

  3   1.9791          4    .7291          5   5.8616

U-STATISTIC              .59745     DEGREES OF FREEDOM    2    1  198
APPROXIMATE F          66.36842     DEGREES OF FREEDOM    2  197.00

F MATRIX - DEGREES OF FREEDOM     2  197

              GROUP
               ACC
GROUP
 UNACC     66.36841

           FUNCTION
               ACC          UNACC
VARIABLE
  1         1.81947        2.99634
  2         2.14663        3.62654

CONSTANT
           -3.66492       -8.98081

              NUMBER OF CASES CLASSIFIED INTO GROUP -
                ACC   UNACC
GROUP
    ACC         78     22
  UNACC         22     78

**************************************************************************

STEP NUMBER            3
VARIABLE ENTERED       5

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1   196

  1  15.2362          2  10.4639          5   5.8616

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1   195

  3    .0645          4   3.4448

U-STATISTIC              .58010     DEGREES OF FREEDOM    3    1  198
APPROXIMATE F          47.29137     DEGREES OF FREEDOM    3  196.00

F MATRIX - DEGREES OF FREEDOM     3  196

              GROUP
               ACC
GROUP
 UNACC     47.29136

           FUNCTION
```

	ACC	UNACC
VARIABLE		
1	1.91947	3.16562
2	1.67710	2.83173
5	.65847	1.11464
CONSTANT	-3.87143	-9.57256

NUMBER OF CASES CLASSIFIED INTO GROUP -

	ACC	UNACC
GROUP		
ACC	80	20
UNACC	20	80

****

STEP NUMBER 4
VARIABLE ENTERED 4

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 195

1 18.0225 2 5.6486 4 3.4448 5 8.6220

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM 1 194

3 1.4660

U-STATISTIC .57003 DEGREES OF FREEDOM 4 1 198
APPROXIMATE F 36.77213 DEGREES OF FREEDOM 4 195.00

F MATRIX - DEGREES OF FREEDOM 4 195

	GROUP ACC
GROUP	
UNACC	36.77212

FUNCTION

	ACC	UNACC
VARIABLE		
1	2.48923	4.21784
2	1.39841	2.31706
4	-.73684	-1.36077
5	1.23894	2.18663
CONSTANT	-3.95552	-9.85936

NUMBER OF CASES CLASSIFIED INTO GROUP -

	ACC	UNACC
GROUP		
ACC	81	19
UNACC	22	78

****

STEP NUMBER 5
VARIABLE ENTERED 3

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM 1 194

1	17.7266	2	4.2594	3	1.4660	4	4.8532
5	8.2265						

U-STATISTIC .56575 DEGREES OF FREEDOM 5 1 198
APPROXIMATE F 29.78121 DEGREES OF FREEDOM 5 194.00

F MATRIX - DEGREES OF FREEDOM 5 194

	GROUP ACC
GROUP UNACC	29.78120

F LEVEL INSUFFICIENT FOR FURTHER COMPUTATION

VARIABLE	FUNCTION ACC	UNACC
1	2.88134	4.95263
2	1.28781	2.10980
3	-.31908	-.59794
4	-1.13178	-2.10086
5	1.74598	3.13679
CONSTANT	-3.98989	-9.90005

GROUP WITH LARGEST PROB.		SQUARE OF DISTANCE FROM AND POSTERIOR PROBABILITY FOR GROUP -			
GROUP ACC		ACC		UNACC	
CASE					
1	ACC	7.703	.662,	9.050	.338,
2	ACC	8.057	.960,	14.426	.040,
3	ACC	4.497	.710,	6.285	.290,
4	ACC	6.164	.876,	10.084	.124,
5	ACC	7.137	.993,	17.018	.007,
6	ACC	12.005	.873,	15.856	.127,
7	UNACC	10.377	.223,	7.876	.777,
8	ACC	9.031	.988,	17.882	.012,
9	ACC	7.705	.934,	13.021	.066,
10	ACC	3.435	.759,	5.732	.241,
11	ACC	2.727	.559,	3.199	.441,
12	ACC	3.553	.843,	6.915	.157,
13	UNACC	6.237	.271,	4.259	.729,
14	ACC	7.876	.817,	10.868	.183,
15	ACC	8.670	.963,	15.172	.037,
16	ACC	2.583	.906,	7.120	.094,
17	ACC	5.289	.622,	6.289	.378,
18	ACC	7.160	.991,	16.523	.009,
19	ACC	2.996	.951,	8.932	.049,
20	ACC	7.233	.505,	7.269	.495,

21	ACC	8,698	.995	19.105	,005
22	ACC	6,943	.912	11.614	,088
23	UNACC	13,145	.225	10.673	,775
24	ACC	2,247	.895	6.531	,105
25	ACC	2,046	.805	4.880	,195
26	ACC	4,305	.987	12.992	,013
27	ACC	5,926	.955	12.032	,045
28	ACC	2,916	.963	9.459	,037
29	ACC	3,482	.691	5.092	,309
30	ACC	6,164	.775	8.636	,225
31	ACC	3,249	.608	4.127	,392
32	ACC	7,684	.907	12.228	,093
33	UNACC	11,587	.025	4.206	,975
34	UNACC	7,842	.408	7.097	,592
35	ACC	4,931	.929	10.076	,071
36	ACC	7,483	.502	7.498	,498
37	UNACC	6,393	.269	4.394	,731
38	ACC	5,529	.849	8.987	,151
39	UNACC	5,035	.268	3.022	,732
40	ACC	5,120	.605	5.977	,395
41	ACC	3,147	.751	5.358	,249
42	ACC	7,913	.943	13.524	,057
43	ACC	7,995	.805	10.826	,195
44	UNACC	5,120	.492	5.056	,508
45	ACC	2,959	.721	4.858	,279
46	ACC	8,553	.993	18.435	,007
47	ACC	7,540	.905	12.041	,095
48	ACC	6,731	.560	7.214	,440
49	ACC	6,876	.796	9.598	,204
50	UNACC	5,281	.273	3.318	,727
51	ACC	7,203	.895	11.495	,105
52	ACC	10,588	.855	14.144	,145
53	ACC	6,587	.983	14.668	,017
54	UNACC	4,793	.235	2.432	,765
55	UNACC	11,768	.292	9.996	,708
56	ACC	4,172	.977	11.627	,023
57	UNACC	7,983	.250	5.781	,750
58	ACC	6,436	.796	9.155	,204
59	ACC	4,040	.958	10.282	,042
60	ACC	5,114	.846	8.527	,154
61	ACC	5,662	.933	10.940	,067
62	ACC	3,908	.890	8.086	,110
63	ACC	8,348	.678	9.839	,322
64	ACC	2,651	.957	8.841	,043
65	UNACC	6,494	.486	6.381	,514
66	ACC	8,885	.512	8.984	,488
67	ACC	4,204	.795	6.916	,205
68	ACC	8,966	.989	17.982	,011
69	ACC	7,136	.579	7.773	,421
70	ACC	8,594	.936	13.974	,064
71	ACC	5,687	.802	8.481	,198
72	ACC	2,788	.793	5.479	,207
73	ACC	7,657	.793	10.341	,207
74	UNACC	9,053	.391	8.170	,609
75	UNACC	5,180	.315	3.630	,685
76	ACC	5,931	.941	11.486	,059
77	ACC	2,593	.875	6.480	,125

78	ACC	4.772	.910,	9.411	.090,
79	UNACC	6.095	.278,	4.185	.722,
80	ACC	3.496	.575,	4.104	.425,
81	ACC	5.527	.960,	11.889	.040,
82	ACC	6.996	.635,	8.100	.365,
83	ACC	2.529	.542,	2.864	.458,
84	ACC	6.977	.717,	8.834	.283,
85	UNACC	21.828	.253,	19.662	.747,
86	ACC	3.669	.839,	6.968	.161,
87	ACC	7.328	.897,	11.661	.103,
88	ACC	2.888	.906,	7.430	.094,
89	ACC	13.101	.882,	17.115	.118,
90	UNACC	4.160	.462,	3.859	.538,
91	ACC	10.238	.948,	16.030	.052,
92	ACC	3.622	.953,	9.660	.047,
93	UNACC	9.537	.476,	9.343	.524,
94	ACC	4.061	.976,	11.509	.024,
95	ACC	1.625	.742,	3.742	.258,
96	ACC	4.312	.909,	8.920	.091,
97	ACC	5.943	.988,	14.753	.012,
98	ACC	4.248	.910,	8.883	.090,
99	ACC	10.305	.993,	20.336	.007,
100	UNACC	6.913	.389,	6.010	.611,

GROUP UNACC CASE			ACC		UNACC
1	UNACC	8.732	.105,	4.448	.895,
2	UNACC	8.501	.254,	6.348	.746,
3	ACC	3.365	.870,	7.172	.130,
4	UNACC	14.319	.248,	12.098	.752,
5	UNACC	18.899	.014,	10.326	.986,
6	ACC	6.136	.801,	8.915	.199,
7	UNACC	8.189	.407,	7.432	.593,
8	UNACC	7.375	.131,	3.599	.869,
9	UNACC	8.380	.057,	2.779	.943,
10	UNACC	6.146	.205,	3.436	.795,
11	UNACC	14.334	.018,	6.284	.982,
12	UNACC	11.089	.176,	8.009	.824,
13	ACC	8.428	.553,	8.857	.447,
14	UNACC	4.088	.317,	2.552	.683,
15	UNACC	10.296	.073,	5.224	.927,
16	ACC	3.669	.838,	6.960	.162,
17	UNACC	3.169	.483,	3.033	.517,
18	UNACC	13.372	.047,	7.336	.953,
19	ACC	4.378	.897,	8.712	.103,
20	UNACC	8.276	.332,	6.874	.668,
21	UNACC	21.418	.014,	12.874	.986,
22	UNACC	4.053	.290,	2.264	.710,
23	UNACC	5.219	.165,	1.980	.835,
24	ACC	1.720	.787,	4.335	.213,
25	ACC	6.050	.504,	6.683	.496,
26	ACC	2.467	.559,	2.937	.441,
27	UNACC	7.887	.097,	3.425	.903,
28	UNACC	9.815	.142,	6.214	.858,
29	UNACC	8.545	.069,	3.351	.931,
30	UNACC	9.503	.318,	7.974	.682,

31	UNACC	24.979	.001	11.556	.999
32	UNACC	14.696	.032	7.878	.968
33	UNACC	5.806	.386	4.879	.614
34	UNACC	13.468	.046	7.411	.954
35	UNACC	14.542	.017	6.471	.983
36	UNACC	8.121	.213	5.507	.787
37	UNACC	13.017	.017	4.932	.983
38	UNACC	10.284	.069	5.068	.931
39	UNACC	6.876	.127	3.016	.873
40	UNACC	8.553	.443	8.091	.557
41	UNACC	10.927	.165	7.686	.835
42	UNACC	11.202	.044	5.066	.956
43	UNACC	7.296	.110	3.123	.890
44	ACC	4.651	.870	8.460	.130
45	UNACC	8.987	.313	7.416	.687
46	UNACC	12.170	.058	6.581	.942
47	UNACC	10.449	.158	7.099	.842
48	UNACC	13.391	.018	5.357	.982
49	UNACC	8.851	.291	7.071	.709
50	UNACC	12.816	.221	10.300	.779
51	ACC	4.771	.732	6.781	.268
52	UNACC	13.222	.015	4.790	.985
53	UNACC	19.587	.019	11.742	.981
54	ACC	2.696	.666	4.079	.334
55	UNACC	16.384	.016	8.110	.984
56	UNACC	9.758	.158	6.405	.842
57	ACC	3.972	.521	4.141	.479
58	UNACC	7.971	.209	5.312	.791
59	UNACC	6.495	.402	5.700	.598
60	UNACC	5.688	.279	3.793	.721
61	UNACC	12.888	.092	8.306	.908
62	ACC	2.604	.515	2.722	.485
63	UNACC	12.517	.043	6.333	.957
64	UNACC	15.050	.048	9.077	.952
65	UNACC	7.449	.157	4.089	.843
66	ACC	5.925	.813	8.869	.187
67	UNACC	12.530	.062	7.096	.938
68	UNACC	9.231	.414	8.538	.586
69	UNACC	9.029	.163	5.751	.837
70	UNACC	6.313	.156	2.931	.844
71	UNACC	10.983	.155	7.595	.845
72	UNACC	15.957	.030	9.001	.970
73	UNACC	12.791	.022	5.169	.978
74	UNACC	6.430	.436	5.914	.564
75	UNACC	4.658	.251	2.473	.749
76	UNACC	6.211	.328	4.777	.672
77	UNACC	14.119	.018	6.137	.982
78	UNACC	8.860	.061	3.396	.939
79	ACC	5.252	.536	5.542	.464
80	UNACC	11.849	.077	6.882	.923
81	UNACC	8.290	.054	2.554	.946
82	UNACC	10.902	.108	6.687	.892
83	UNACC	30.036	.016	21.798	.984
84	UNACC	6.569	.200	3.795	.800
85	UNACC	9.297	.295	7.558	.705
86	UNACC	4.348	.318	2.818	.682
87	UNACC	15.079	.263	13.022	.737

88	UNACC	10.653	.040,	4.279	.960,
89	UNACC	10.870	.465,	10.589	.535,
90	UNACC	3.611	.496,	3.578	.504,
91	UNACC	16.069	.042,	9.803	.958,
92	ACC	2.710	.666,	4.086	.334,
93	UNACC	5.602	.122,	1.646	.878,
94	UNACC	5.917	.325,	4.453	.675,
95	ACC	3.327	.797,	6.065	.203,
96	UNACC	5.707	.328,	4.270	.672,
97	ACC	6.098	.879,	10.057	.121,
98	UNACC	13.635	.030,	6.659	.970,
99	ACC	6.236	.604,	7.084	.396,
100	UNACC	10.329	.049,	4.377	.951,

NUMBER OF CASES CLASSIFIED INTO GROUP -

GROUP	ACC	UNACC
ACC	80	20
UNACC	19	81

SUMMARY TABLE

STEP NUMBER	VARIABLE ENTERED	VARIABLE REMOVED	F VALUE TO ENTER OR REMOVE	NUMBER OF VARIABLES INCLUDED	U-STATISTIC
1	2		11.2828	1	.6402
2	1		14.0945	2	.5974
3	5		5.8616	3	.5801
4	4		3.4448	4	.5700
5	3		1.4660	5	.5658

BIBLIOGRAPHY

1. Afifi, A. A. and Azen, S. P., Statistical Analysis, A Computer Oriented Approach, Academic Press, New York, 1972.

2. Anderberg, Michael R., Cluster Analysis for Application, Academic Press, New York, 1973.

3. Biomedical Computer Programs, Health Sciences Computing Facility, Department of Biomathematics, University of California Press, Los Angeles, 1973.

4. Bolch, B. W. and Huang, C. J., Multivariate Statistical Methods for Business and Economics, Prentice Hall, Inc., New Jersey, 1974.

5. Chernoff, Herman, "Metric Considerations in Cluster Analysis," In Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, pp. 621-629, University of California Press, Berkeley, 1970, AD714810.

6. Comrey, Andrew L., A First Course in Factor Analysis, Academic Press, New York, 1973.

7. Cole, A. J., Numerical Taxonomy, Academic Press, New York, 1969.

8. Collons, Rodger D., An Algorithm for the Coefficient of Concordance, Research Paper Number 41, Georgia State College, Atlanta, Feb. 1968.

9. Cooley, William W. and Lohnes, Paul R., Multivariate Data Analysis, John Wiley and Sons, Inc., New York, 1971.

10. Division Command Post Test (FM 286), Draft Detailed Plan for Execution, MASSTER, Fort Hood, Texas, November 1974.

11. Division Command Post Test Report (FM 286), MASSTER, Fort Hood, Texas, March 1975.

12. Downie, N. M. and Heath, R. W., Basic Statistical Methods, Harper and Row, New York, 1970.

13. Dunnette, M. D. and Kirchner, W. K., Psychology Applied to Industry, Meredith Corporation, New York, 1965.

14. Edwards, Allen 1., Statistical Methods, Harper and Row, New York, 1973.

15. Edwards, Allen L., Statistical Analysis, Holt, Rinehart, and Winston, Inc., New York, 1974.

16. Everitt, Brian, Cluster Analysis, John Wiley and Sons, New York, 1974.

17. Fisher, R. A., "The Precision of Discriminant Functions," Annals of Eugenics, Vol. X, Pt. IV, pp, 422-429, 1940.

18. Fisher, R. A., "The Statistical Utilization of Multiple Measurements," Annals of Eugenics, Vol. VIII, Pt. IV, pp. 376-386, 1938.

19. Fisher, R. A., "The Use of Multiple Measurements in Taxonomic Problems," Annals of Eugenics, Vol. VII, PT.II, pp. 179-188, 1936.

20. Fisher, Lloyd and Van Ness, John, "Admissible Clustering Procedures," Biometrika, Vol. 58, No. 1, 1971.

21. Fisher, Lloyd and Van Ness, John, "Admissible Discriminant Analysis," Journal of the American Statistical Association, Vol. 68, No. 343, pp. 603-607, September 1973.

22. Fishman, George S., Concepts and Methods in Discrete Event Digital Simulation, John Wiley and Sons, New York, 1973.

23. Friedman, H. P. and Rubin, J., "On Some Invariant Criteria for Grouping Data," Journal of the American Statistical Association, Vol. 62, pp. 1159-1178, 1967.

24. Green, Paul E., Grank, Ronald E., and Robinson, Patrick J., "Cluster Analysis in Test Market Selection," Management Science, Vol. 13, No. 8, April 1967.

25. Hall, D. J., Duda, R. O., and Huffman, D. A., "Development of New Pattern-Recognition Methods," F33615-71-C1894, Stanford Research Institute, Menlo Park, California, November 1973, AD 772614.

26. Hillier, F. S., The Evaluation of Risky Interrelated Investments, North-Holland, 1969.

27. Hines, William W. and Montgomery, Douglas C., Probability and Statistics in Engineering and Management Science, The Ronald Press Company, New York, 1972.

28. Homer, Eugene D., "A Generalized Model for Analyzing Management Information Systems, "Management Science, Vol. 8, 1962.

29. Integrated Battlefield Control System Division Level System Definition, Second Refinement (Final Study), MASSTER, Fort Hood, Texas, February 1973.

30. Kabe, D. G. and Gupta, R. P., Multivariate Statistical Inference, Americal Elaevier Publishing Co., New York, 1973.

31. Lieberman, Irving J., "A Mathematical Model for Integrated Business Systems," An address presented at the Second National Meeting of the Institute of Management Sciences, New York, October 20-21, 1955.

32. Ling, Robert F., "A Probability Theory of Cluster Analysis," Journal of the American Statistical Association, Vol. 68. No. 341, March 1973.

33. Malkovich, J. F., and Afifi, A. A., "On Tests for Multivariate Normality," Journal of the American Statistical Association, Vol. 68, No. 341, pp. 176-179.

34. MASSTER-Test Officer's Planning Manual, Second Edition, Headquarters, Modern Army Selected Systems Test, Evaluation, and Review, January 1975.

35. Moore, John R. Jr., and Baker, Norman R., An Analytical Approach to Scoring Model Design, Graduate School of Business, Stanford University, 1969.

36. Morrison, Donald G., "Measurement Problems in Cluster Analysis," Management Science, Vol. 13, No. 12, August 1967.

37. Morse, Philip M., "Optimal Linear Ordering of Information Items," Operations Research, July-August 1972.

38. Operational Test Methodology Guide, Volume II, Techniques and Guidelines, Operational Test and Evaluation Agency, Fort Belvoir, Virginia, January 1974.

39. Overall, John E. and Klett, C. James, Applied Multivariate Analysis, McGraw-Hill, New York, 1972.

40. Rao, M. R., "Cluster Analysis and Mathematical Programming," Journal of the American Statistical Association, Vol. 66, No. 335, September 1971.

41. Williamsen, Eleanor W., Understanding Statistical Reasoning, W. H. Freeman and Company, San Francisco, 1974.